

## Лабораторная работа №02: Организация данных и системный каталог

**Цель работы:** Всестороннее изучение логической и физической структуры хранения данных в PostgreSQL. Получение практических навыков управления базами данных, схемами, табличными пространствами. Глубокое освоение работы с системным каталогом для извлечения метаданных. Исследование низкоуровневых аспектов хранения, включая TOAST.

### Стек технологий:

- **ОС:** Xubuntu 24.04 64-bit (предустановлена в виртуальной машине)
- **СУБД:** PostgreSQL 16
- **Утилиты:** `psql`
- **Ключевые понятия:** База данных, схема, табличное пространство, системный каталог, `pg_class`, `pg_tables`, `pg_database`, `pg_tablespace`, `search_path`, `temp_buffers`, TOAST, нежурналируемые таблицы, стратегии хранения

### Теоретическая часть (краткое содержание):

- **Логическая структура:** Иерархия: Кластер БД -> Базы данных -> Схемы -> Объекты (таблицы, представления и т.д.).
- **Физическая структура:** Табличные пространства связывают логические объекты с физическим расположением файлов на диске. Данные хранятся в файлах.
- **Системный каталог:** Набор системных таблиц (`pg_class`, `pg_database`, `pg_namespace`) и представлений (`pg_tables`, `pg_views`), содержащих метаданные.
- **Низкоуровневое хранение:** Механизм TOAST для больших данных, нежурналируемые таблицы, стратегии хранения столбцов.

### Задание на практическую реализацию:

#### Модуль 1: Базы данных и схемы

1. **Создание и проверка БД:** Создайте новую базу данных `lab02_db`. Проверьте ее начальный размер с помощью `pg_database_size('lab02_db')`.
2. **Работа со схемами:** Подключитесь к `lab02_db`. Создайте две схемы: `app` и схему с именем вашего пользователя ОС (напр., `student`). В каждой схеме создайте по одной таблице и вставьте в них данные.
3. **Контроль размера:** Снова проверьте размер базы данных. Объясните его изменение.
4. **Управление путем поиска:** Настройте параметр `search_path` для текущего сеанса так, чтобы при обращении по неполному имени приоритет имела ваша пользовательская схема, а затем схема `app`. Продемонстрируйте работу, обратившись к таблицам без указания схемы.
5. **Практика+ (Настройка параметра БД):** Для базы `lab02_db` установите значение параметра `temp_buffers` так, чтобы в каждом новом сеансе, подключенном к этой БД, оно было в 4 раза больше значения по умолчанию. Проверьте работу.

#### Модуль 2: Системный каталог

1. **Исследование `pg_class`:** Получите описание системной таблицы `pg_class` (команда `\d pg_class`).

2. **Исследование `pg_tables`:** Получите подробное описание представления `pg_tables` (команда `\d+ pg_tables`). Объясните разницу между таблицей и представлением.
3. **Временная таблица и список схем:** В базе `lab02_db` создайте временную таблицу. Получите полный список всех схем в этой БД, включая системные (`pg_catalog`, `information_schema`). Объясните наличие временной схемы.
4. **Представления `information_schema`:** Получите список всех представлений в схеме `information_schema`.
5. **Анализ метакоманды:** Выполните в `psql` команду `\d+ pg_views`. Изучите вывод и объясните, какие запросы к системному каталогу скрыты за этой командой.

### Модуль 3: Табличные пространства

1. **Создание Tablespace:** Создайте каталог в файловой системе (напр., `/home/student/mytablespace`). Создайте новое табличное пространство `lab02_ts`, указывающее на этот каталог.
2. **Tablespace по умолчанию:** Измените табличное пространство по умолчанию для базы данных `template1` на `lab02_ts`. Объясните цель этого действия.
3. **Наследование свойства:** Создайте новую базу данных `lab02_db_new`. Проверьте ее табличное пространство по умолчанию. Объясните результат.
4. **Символическая ссылка:** Найдите в каталоге `PGDATA/pg_tblspc/` символическую ссылку, соответствующую `lab02_ts`. Куда она ведет?
5. **Удаление Tablespace:** Удалите табличное пространство `lab02_ts` с опцией `CASCADE`. Объясните необходимость использования `CASCADE`.
6. **Практика+ (Параметр Tablespace):** Установите параметр `random_page_cost` в значение `1.1` для табличного пространства `pg_default`.

### Модуль 4: Низкий уровень

1. **Нежурналируемая таблица:** Создайте нежурналируемую таблицу в пользовательском табличном пространстве (создайте временное для задания). Убедитесь, что для таблицы существует слой `init` (файл с суффиксом `_init`). Удалите табличное пространство.
2. **Стратегии хранения TOAST:**
  - Создайте таблицу со столбцом типа `text`.
  - Определите стратегию хранения по умолчанию для этого столбца.
  - Измените стратегию хранения на `external`.
  - Вставьте в таблицу короткую (менее 2 КБ) и длинную (более 2 КБ) строки.
  - Проверьте, попали ли строки в TOAST-таблицу, выполнив запрос к `pg_toast.pg_toast_<oid_таблицы>`. Объясните результат.
3. **Практика+ (Анализ размера БД):** Сравните размер базы `lab02_db`, возвращаемый `pg_database_size`, с суммой размеров всех ее пользовательских таблиц (`pg_total_relation_size`). Объясните расхождение.
4. **Практика+ (Методы сжатия TOAST):** Проверьте средствами SQL, был ли PostgreSQL скомпилирован с поддержкой методов сжатия `pglz` и `lz4`.
5. **Практика+ (Сравнение сжатия):** Создайте текстовый файл >10 МБ. Загрузите его содержимое в таблицу с текстовым полем трижды: 1) без сжатия (`storage = external`), 2) сжатие `pglz`, 3) сжатие `lz4`. Сравните размер таблицы и время загрузки для каждого варианта.

### Требования к оформлению и отчету:

1. **Полнота:** Должны быть выполнены ВСЕ задания из всех модулей.
2. **Скрипты:** Предоставить полную последовательность выполненных SQL-команд и команд ОС.
3. **Выводы:** Приложить вывод команд, подтверждающих каждое действие.
4. **Отчет:** Подробный отчет с ответами на все вопросы ("объясните", "почему", "проверьте", "сравните") из формулировок заданий. Отчет должен демонстрировать понимание не только команд, но и принципов работы СУБД.

#### Критерии оценки:

- **Удовлетворительно:** Полностью выполнены Модули 1 и 2. Даны ответы на основные вопросы.
- **Хорошо:** Дополнительно полностью выполнен Модуль 3. Отчет содержит развернутые объяснения по всем пунктам этих модулей.
- **Отлично:** Полностью выполнены ВСЕ модули (1, 2, 3, 4), включая все задания "Практика+". Отчет демонстрирует глубокое системное понимание устройства СУБД.

**Рекомендуемая литература:** (Список без изменений, как в предыдущем варианте)

#### Рекомендуемая литература:

1. Курс "Администрирование PostgreSQL 16. Базовый курс":  
<https://postgrespro.ru/education/courses/DBA1>
2. Документация: CREATE DATABASE: <https://postgrespro.ru/docs/postgresql/16/sql-createdatabase>
3. Документация: CREATE SCHEMA: <https://postgrespro.ru/docs/postgresql/16/sql-createschema>
4. Документация: CREATE TABLESPACE: <https://postgrespro.ru/docs/postgresql/16/sql-createtablespace>
5. Документация: Системный каталог: <https://postgrespro.ru/docs/postgresql/16/catalog-pg-class>
6. Документация: Параметры конфигурации: <https://postgrespro.ru/docs/postgresql/16/runtime-config>
7. Книги по PostgreSQL: <https://postgrespro.ru/education/books>