

Выпускная квалификационная работа  
на соискание академической степени магистра  
прикладной математики и информатики  
по направлению «Прикладная математика и информатика»

## Интеграция параллелизма в СУБД PostgreSQL<sup>1</sup>

К.С. Пан

**Рецензент:**

доктор физ.-мат. наук, профессор  
Л.Б. Соколинский

**Научный руководитель:**

канд. физ.-мат. наук, доцент  
М.Л. Цымблер

25 июня 2011

---

<sup>1</sup>Работа выполнена при финансовой поддержке РФФИ (проект

# Актуальность

- ▶ Задачи со **сверхбольшими данными** становятся популярными.
- ▶ Параллельные СУБД **дороги и сложны** в использовании.
- ▶ Свободные СУБД не параллельны, но имеют **открытый код**.

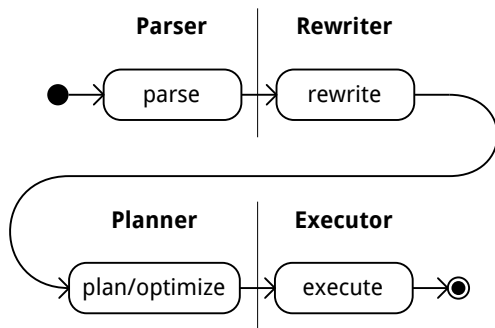
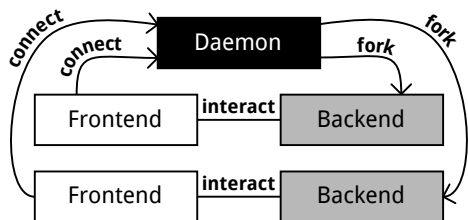
## Цель и задачи

**Цель** — внедрить поддержку фрагментного параллелизма в СУБД PostgreSQL.

### Задачи:

1. Изучить архитектуру СУБД PostgreSQL.
2. Разработать архитектуру и принципы реализации параллельной СУБД PargreSQL на основе внедрения параллелизма в СУБД PostgreSQL.
3. Выполнить проектирование и реализацию основных подсистем параллельной СУБД PargreSQL.
4. Разработать набор тестов и провести эксперименты для исследования эффективности.

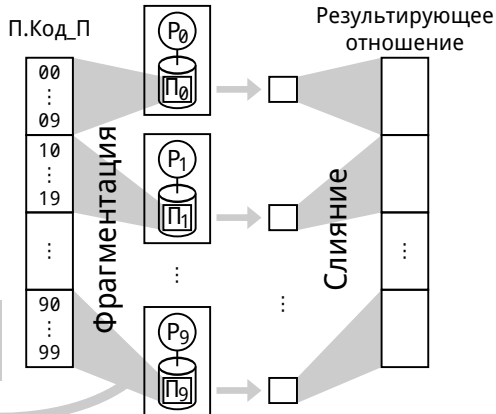
# Архитектура СУБД PostgreSQL



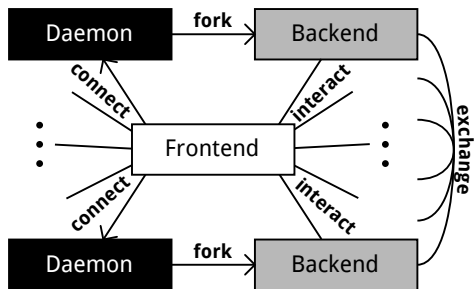
# Фрагментный параллелизм

$$\Pi_i = \{t | t \in \Pi, \phi(t) = i\}$$
$$i = 0, \dots, 9$$

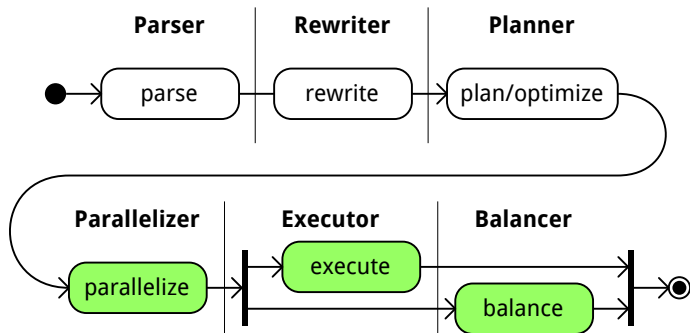
Функция фрагментации  
 $\phi(t) = (t.\text{Код\_}\Pi \text{ div } 10) \bmod 10$



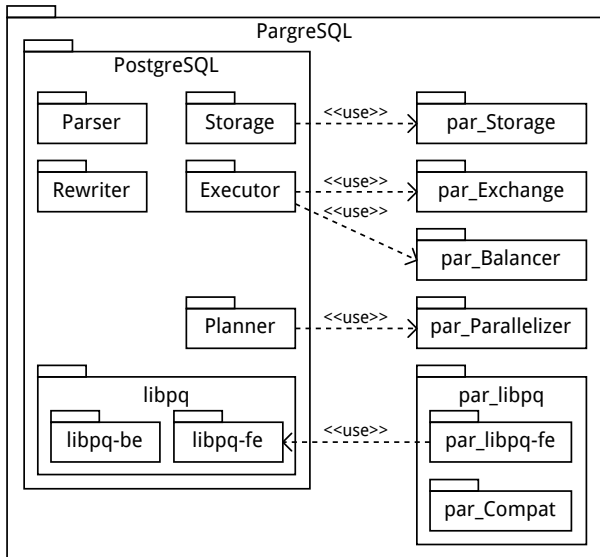
# Клиент-серверная модель СУБД PostgreSQL



# Исполнение запросов в СУБД PostgreSQL



# Архитектура СУБД PargreSQL

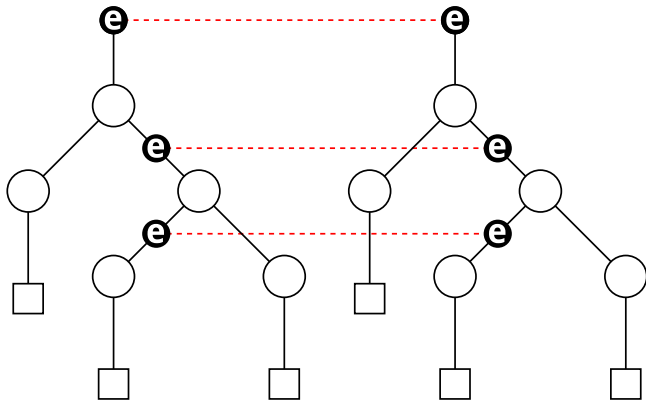




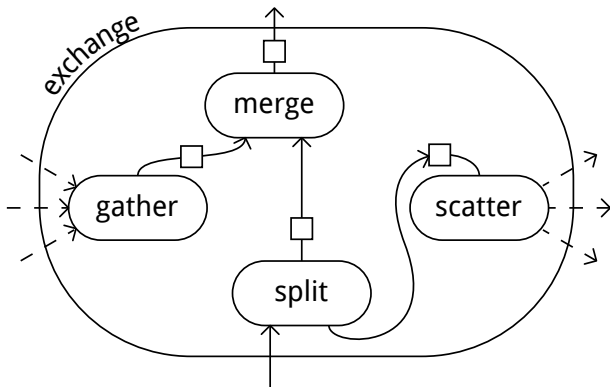
# Оператор Exchange

Выч. узел 0

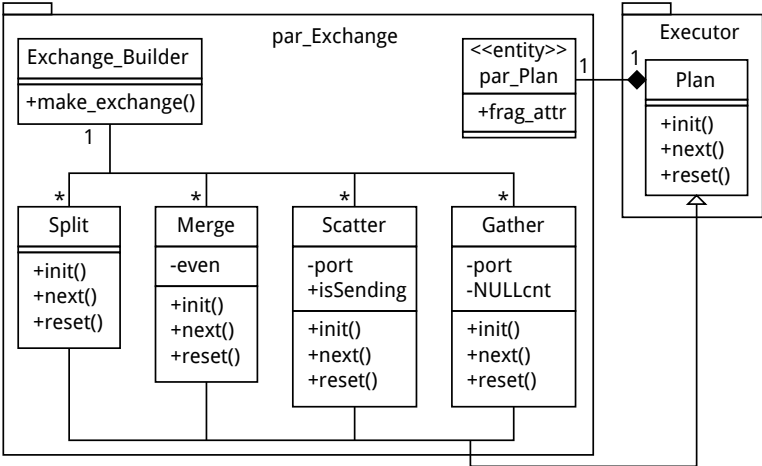
Выч. узел 1



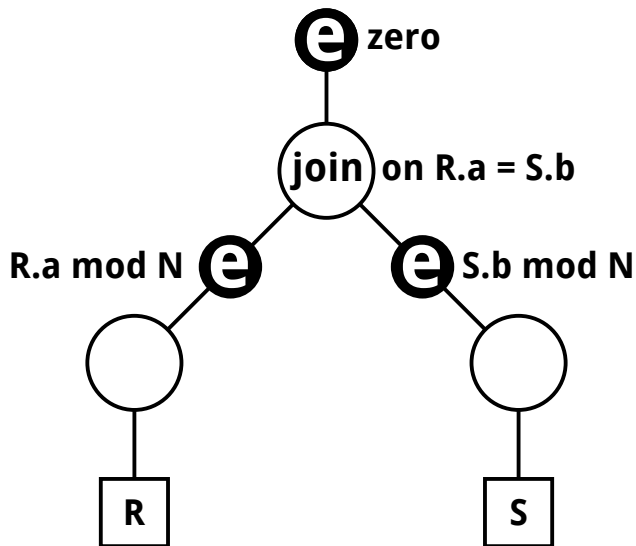
# Архитектура оператора Exchange



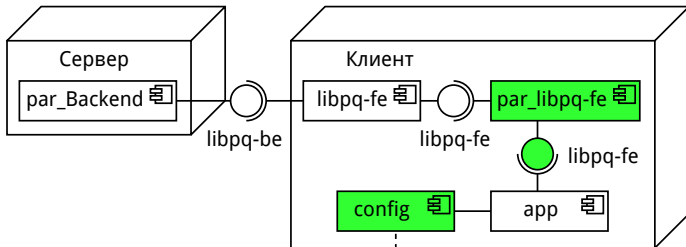
# Реализация Exchange



# Параллелизатор

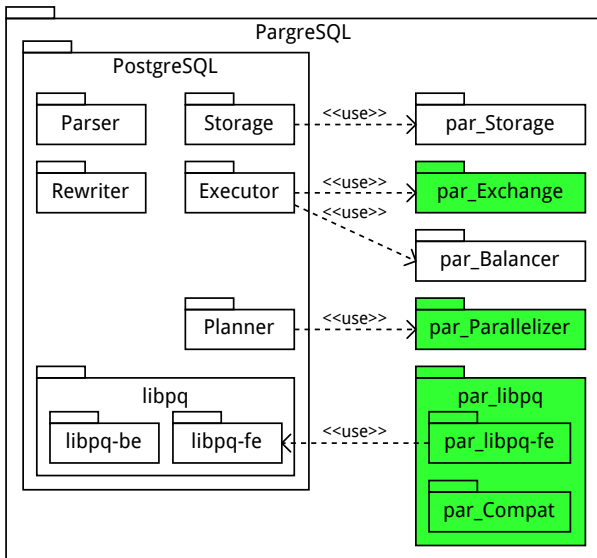


# Изменения на стороне клиента

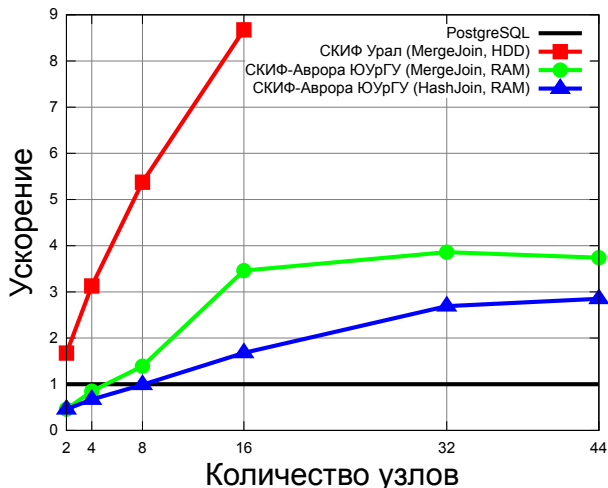


```
<pargresql_config>
  <nodes>
    <node addr="10.0.0.40" port="5432"/>
    <node addr="10.0.0.41" port="5432"/>
  </nodes>
  <partitioning>
    <relation name="T" partattr="f" partfunc="% 2"/>
    <relation name="P" partattr="x" partfunc="/ 2 % 2"/>
  </partitioning>
</pargresql_config>
```

# Разработанные подсистемы СУБД PargreSQL



# Вычислительные эксперименты



## Платформы:

Кластер СКИФ Урал  
2 x Intel Xeon E5472 на узел  
сеть InfiniBand 20 Гбит/с

Кластер СКИФ-Аврора ЮУрГУ  
2 x Intel Xeon X5680 на узел  
сеть 3D-Тор 60 Гбит/с

## Данные:

$S = 10^8$  кортежей  
 $R = 10^7$  кортежей  
итого 2 ГБ

## Запрос:

```
select *  
from R join S  
on R.b = S.a  
where R.a % 10007 = 0
```

## Публикации

1. *Pan C.* Development of a Parallel DBMS on the Basis of PostgreSQL // Proceedings of the Seventh Spring Researchers' Colloquium on Databases and Information Systems (SYRCoDIS'2011). Moscow: Moscow State University, 2011. P. 57–61.
2. *Пан К.С., Цымблер М.Л.* Архитектура и принципы реализации параллельной СУБД PargreSQL // Параллельные вычислительные технологии (ПаВТ'2011): труды международной научной конференции (Москва, 28 марта – 1 апреля 2011 г.). Челябинск: Издательский центр ЮУрГУ, 2011. С. 577–584.
3. *Пан К.С., Цымблер М.Л.* Проект PargreSQL: разработка параллельной СУБД на основе свободной СУБД PostgreSQL // Научный сервис в сети Интернет: суперкомпьютерные центры и задачи: Труды международной научной конференции (Новороссийск, 20–25 сентября 2010 г.). М.: Изд-во МГУ, 2010. С. 308–313.



# Апробации

1. The Seventh Spring Researchers Colloquium on Databases and Information Systems, SYRCoDIS'2011 (June 2–3, 2011, Moscow, Russia).
2. Параллельные вычислительные технологии (ПаВТ'2011). Международная научная конференция (Москва, 28 марта – 1 апреля 2011 г.).
3. "Научный сервис в сети Интернет: суперкомпьютерные центры и задачи": Международная научная конференция (Новороссийск, 20–25 сентября 2010 г.).

# Гранты

- ▶ Грант РФФИ (проект 09-07-00241-а)
- ▶ Конкурс администрации г. Челябинска «Наука и бизнес»

# Основные результаты

1. Разработана архитектура параллельной СУБД PargreSQL на основе свободной последовательной СУБД PostgreSQL.
2. Разработаны интерфейсы и выполнена реализация подсистем:
  - ▶ подсистема тиражирования запросов;
  - ▶ оператор обмена;
  - ▶ параллелизатор плана запроса;
3. Проведены эксперименты на вычислительном кластере СКИФ Урал, результаты которых показывают эффективность разработанной СУБД.