

Международная научная конференция
«Параллельные вычислительные технологии 2011»
(28 марта – 1 апреля 2011 г., Москва)

Архитектура и принципы реализации параллельной СУБД PargreSQL

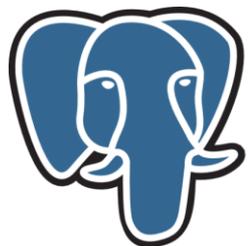
М.Л. Цымблер, К.С. Пан

Южно-Уральский государственный университет (Челябинск)

Работа выполнена при финансовой поддержке РФФИ (проект 09-07-00241-а).

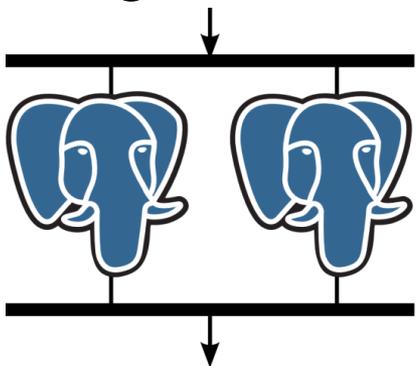
PostgreSQL & PargreSQL

- PostgreSQL
 - Open-source объектно-реляционная СУБД
 - Linux, Windows, Mac OS, Solaris, FreeBSD, QNX
 - SQL:2003
 - Макс размеры: таблица 32 Тб, колонка - 1 Гб



PostgreSQL

- PargreSQL = PostgreSQL + фрагментный параллелизм

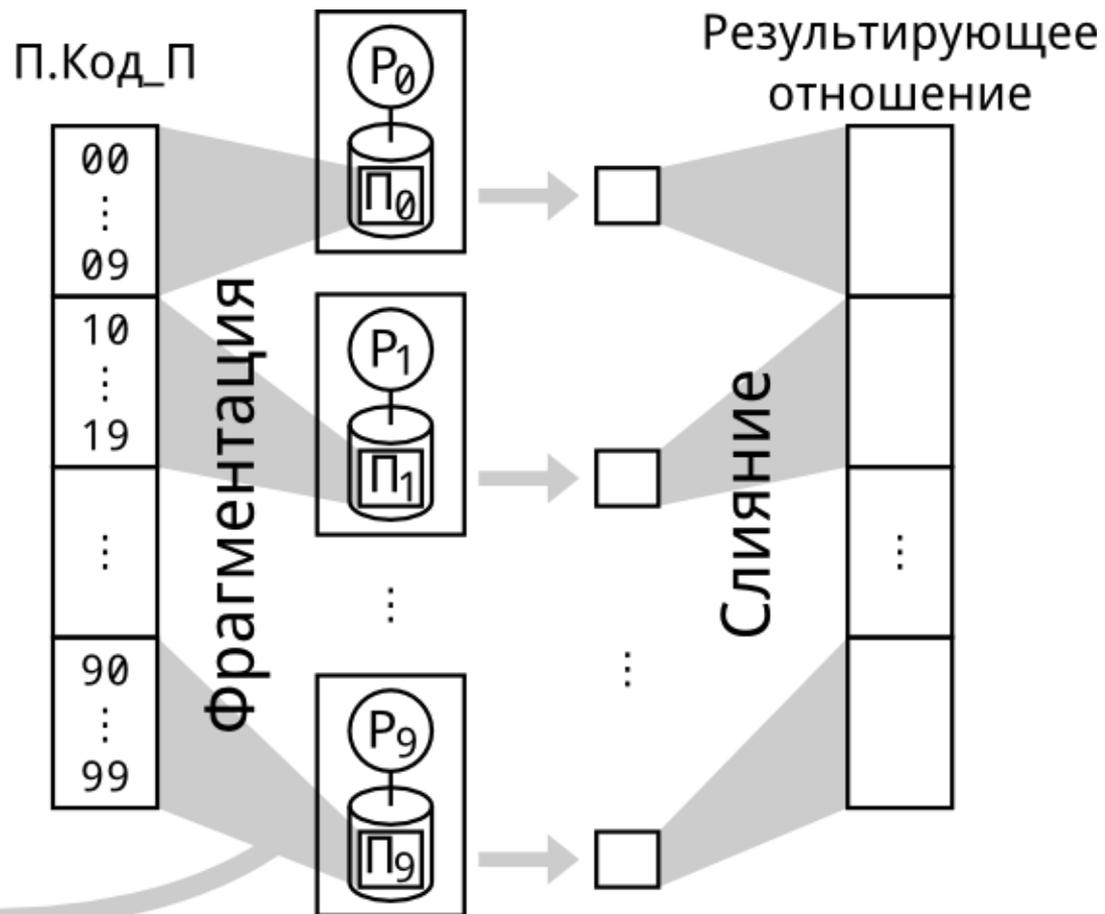


PargreSQL

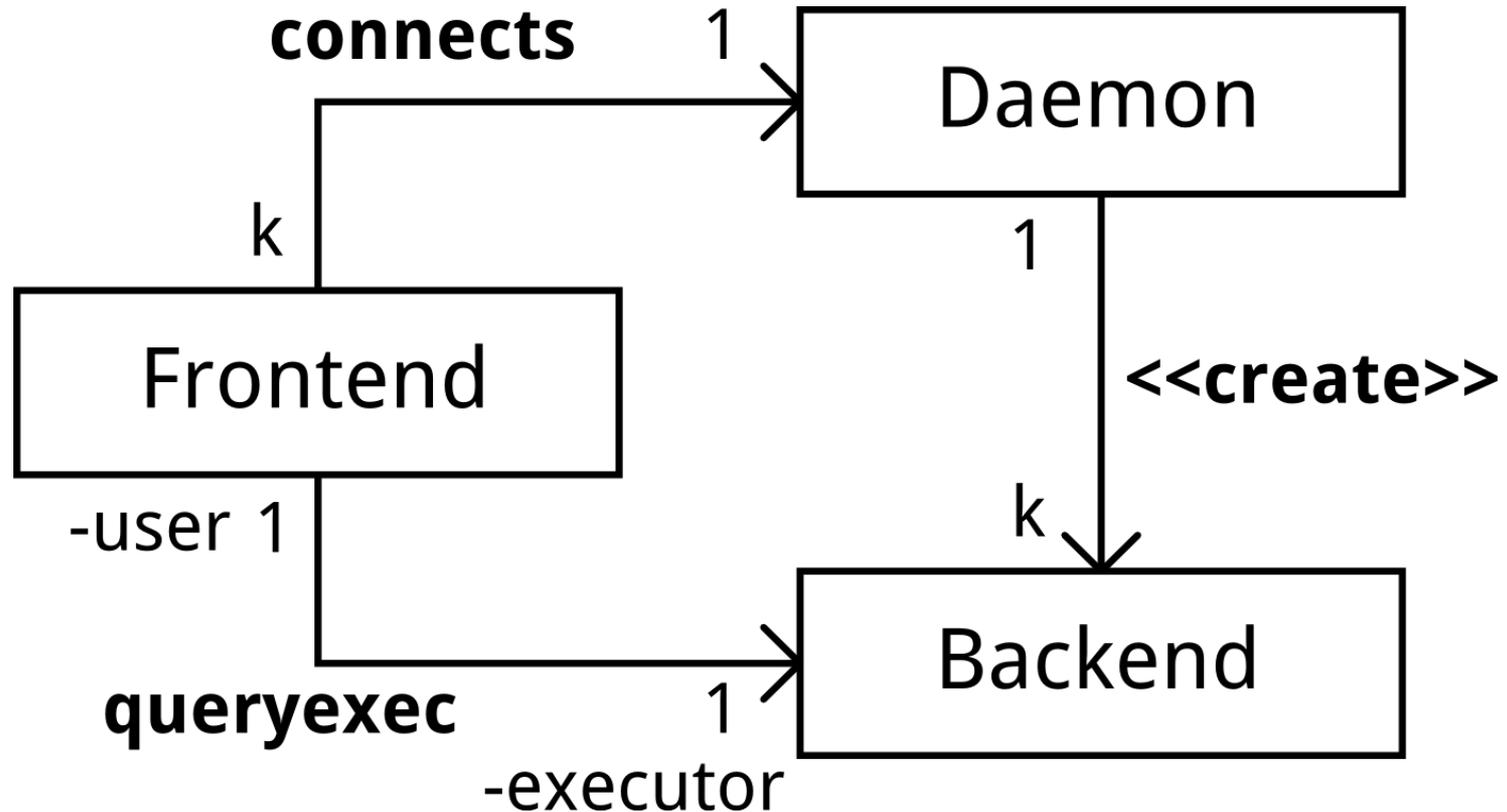
Фрагментный параллелизм

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

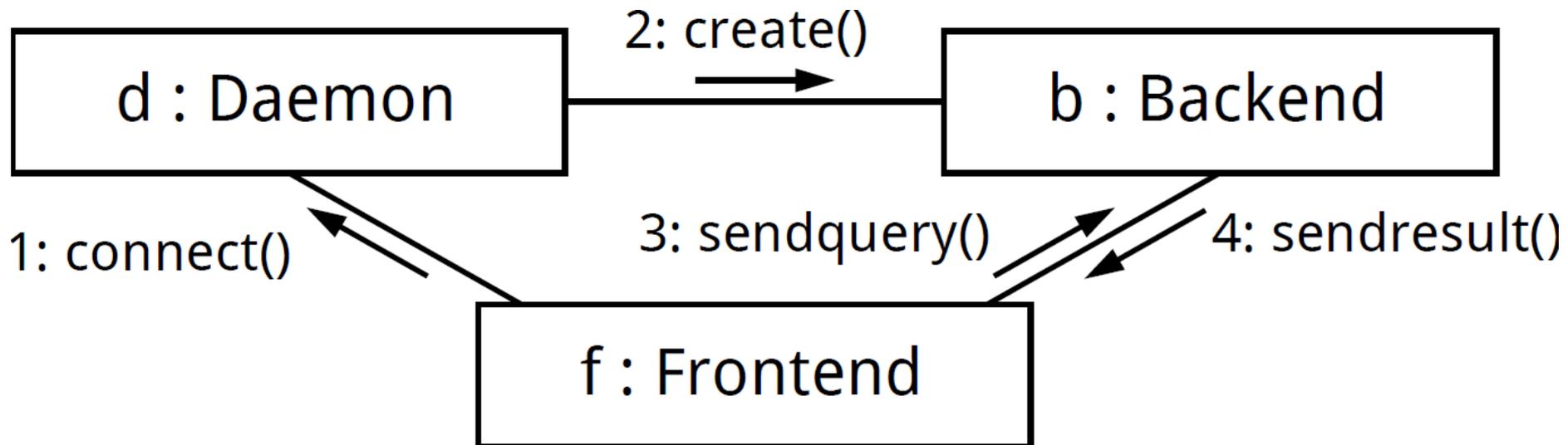
Функция фрагментации
 $\phi(t) = (t.\text{Код_}\Pi \text{ div } 10) \bmod 10$



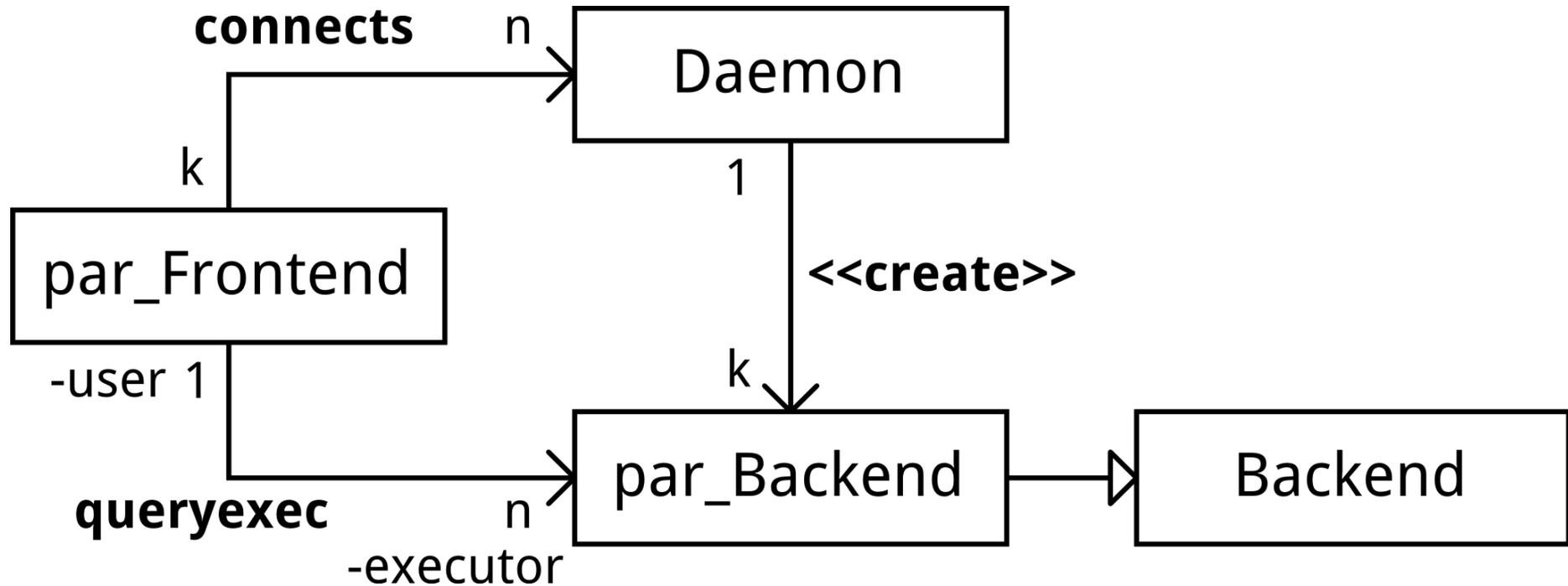
Процессы СУБД: PostgreSQL



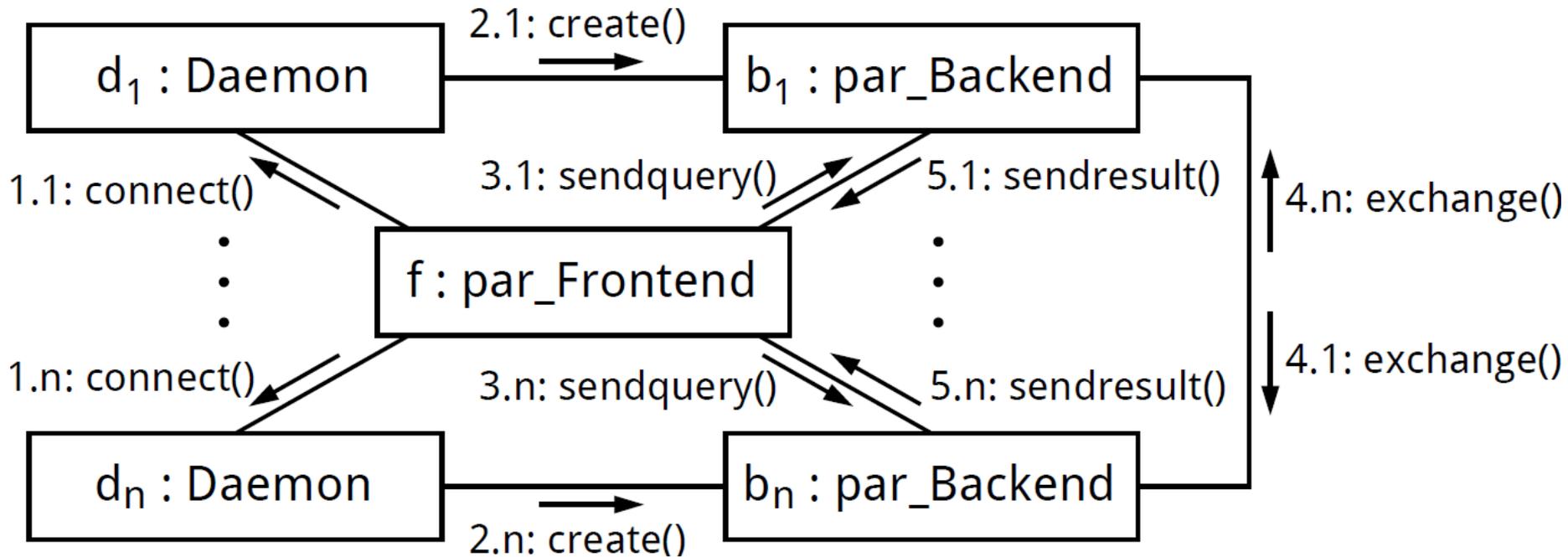
Процессы СУБД: PostgreSQL



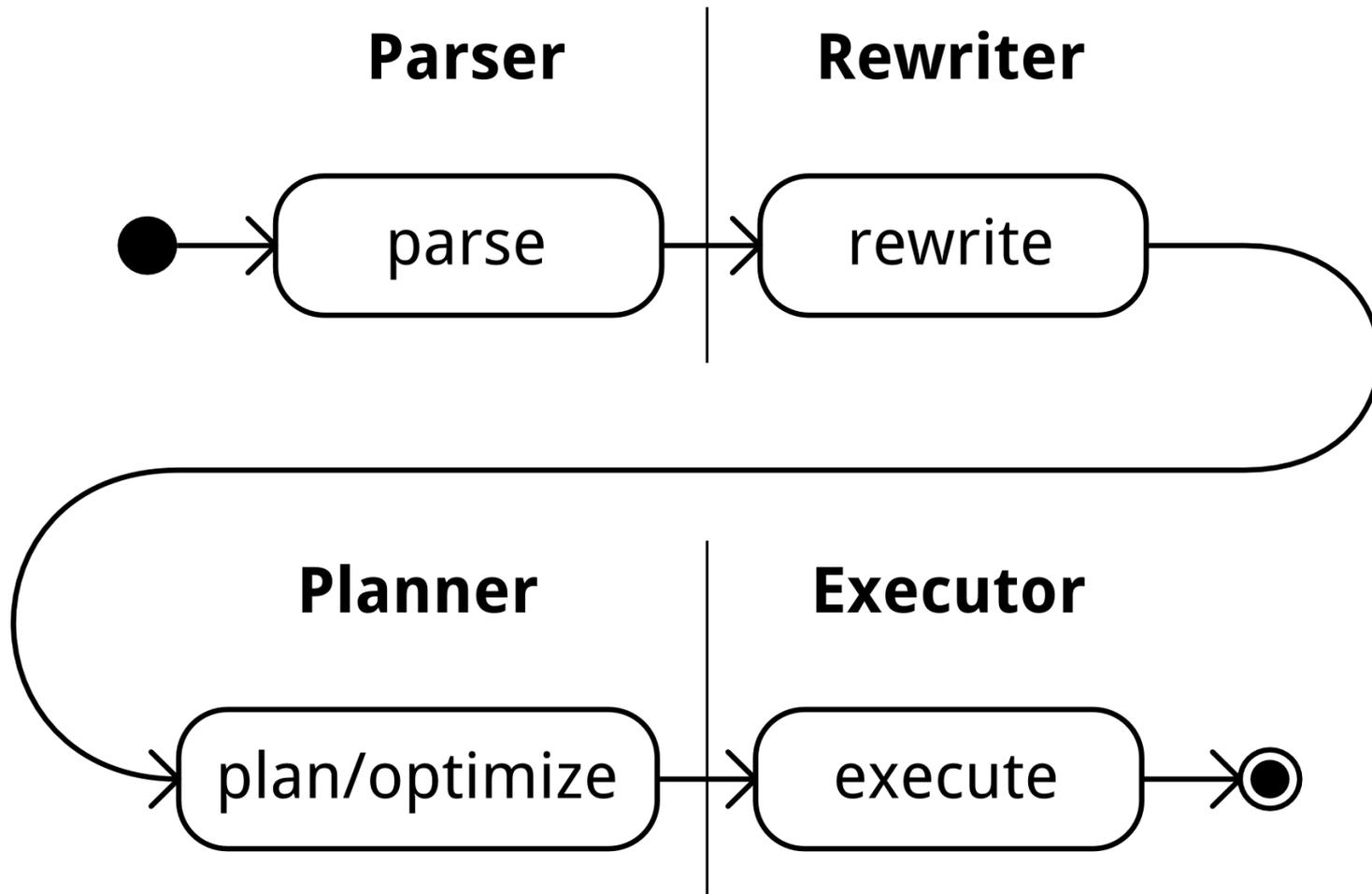
Процессы СУБД: PargreSQL



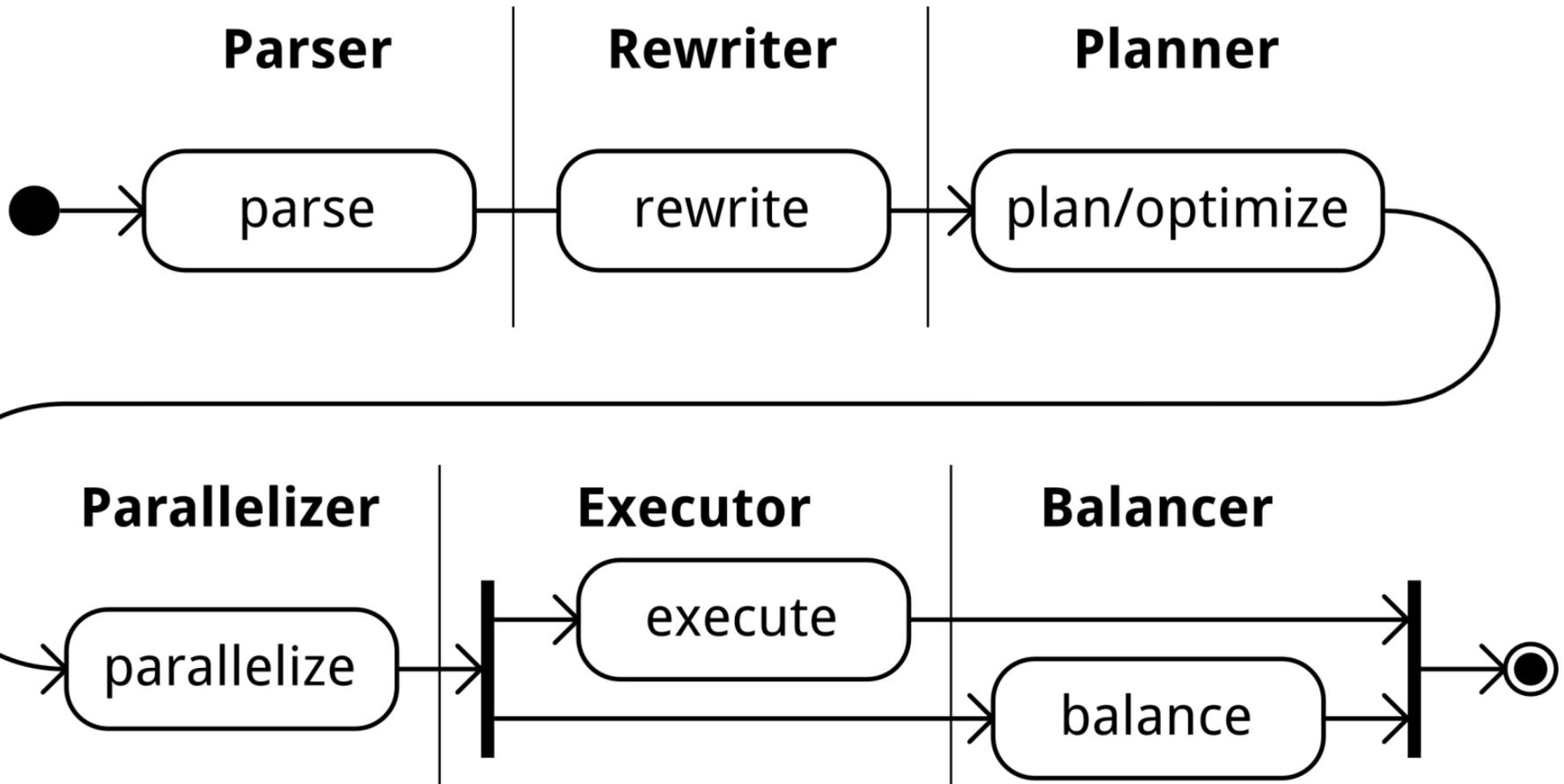
Процессы СУБД: PargreSQL



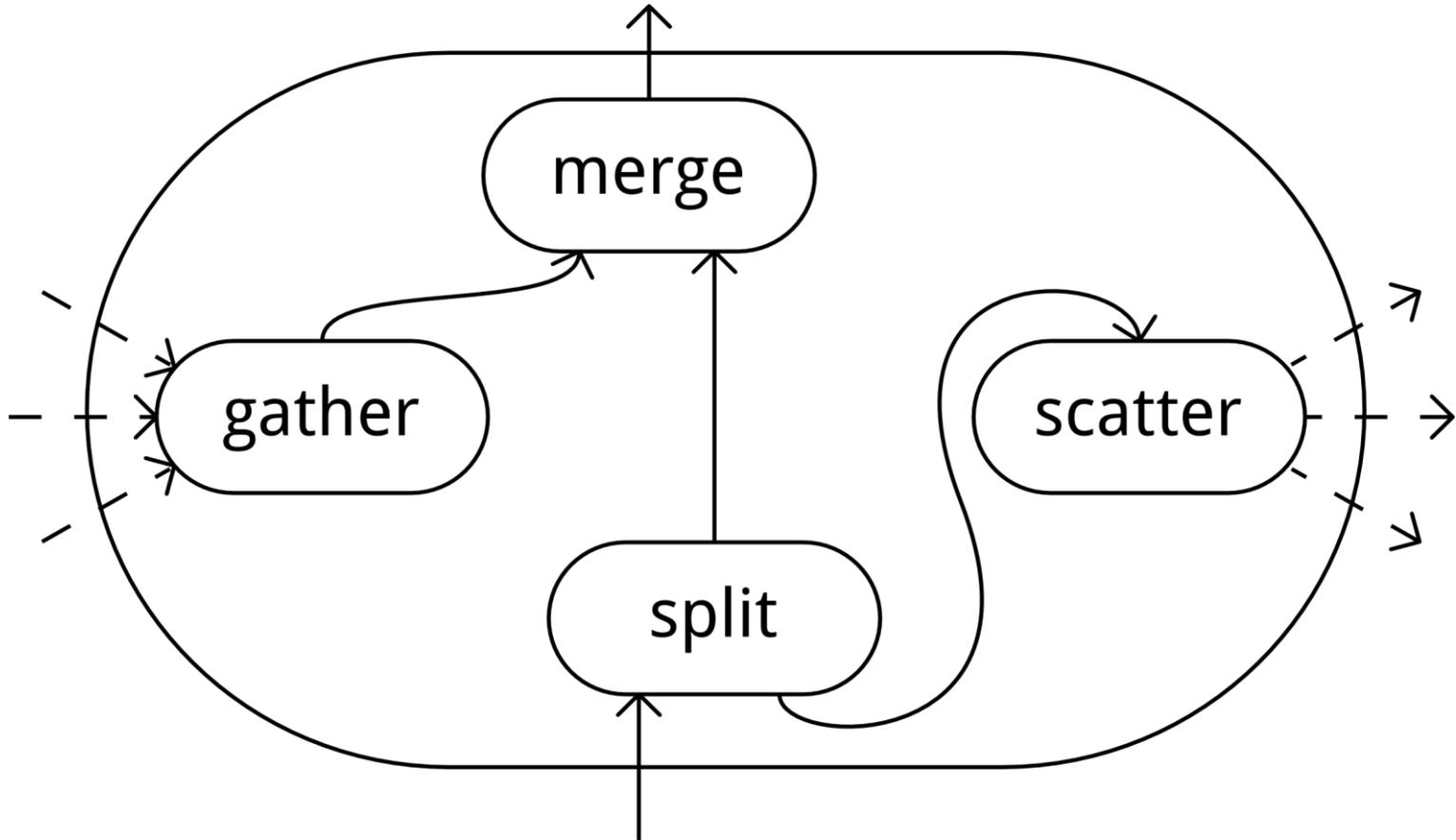
Обработка запроса: PostgreSQL



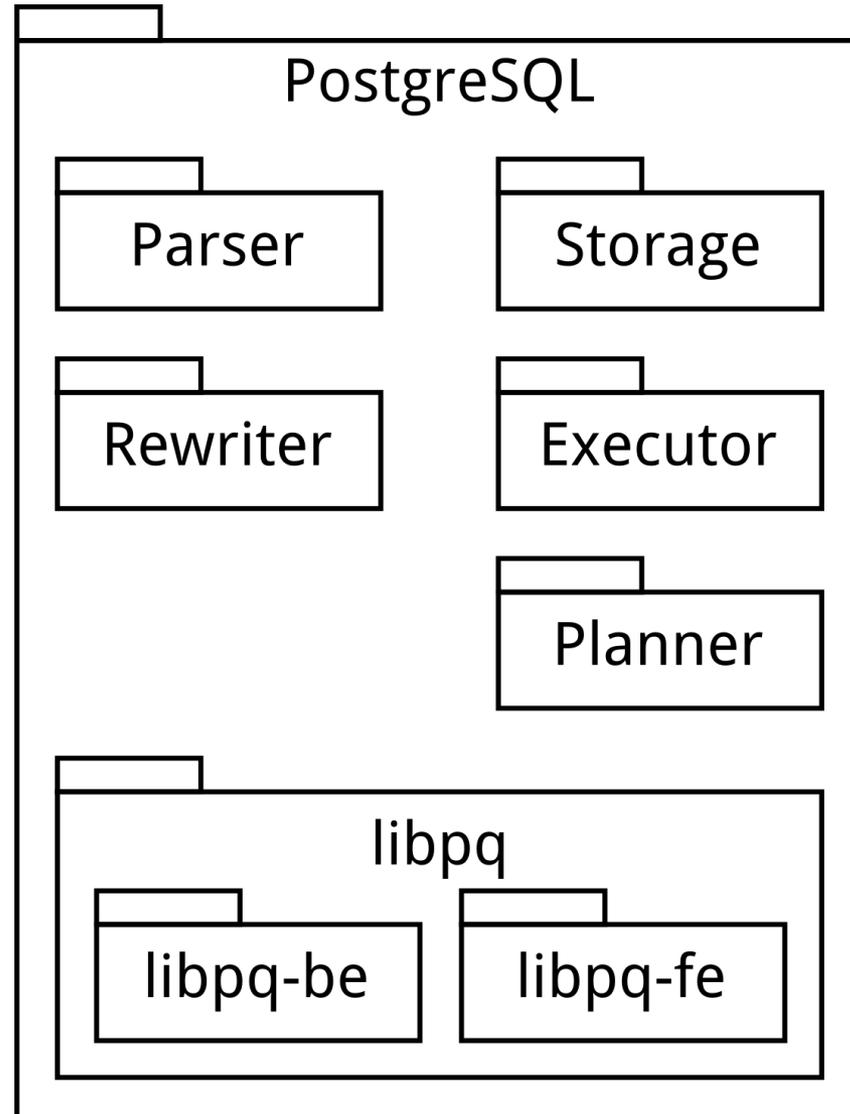
Обработка запроса: PargreSQL



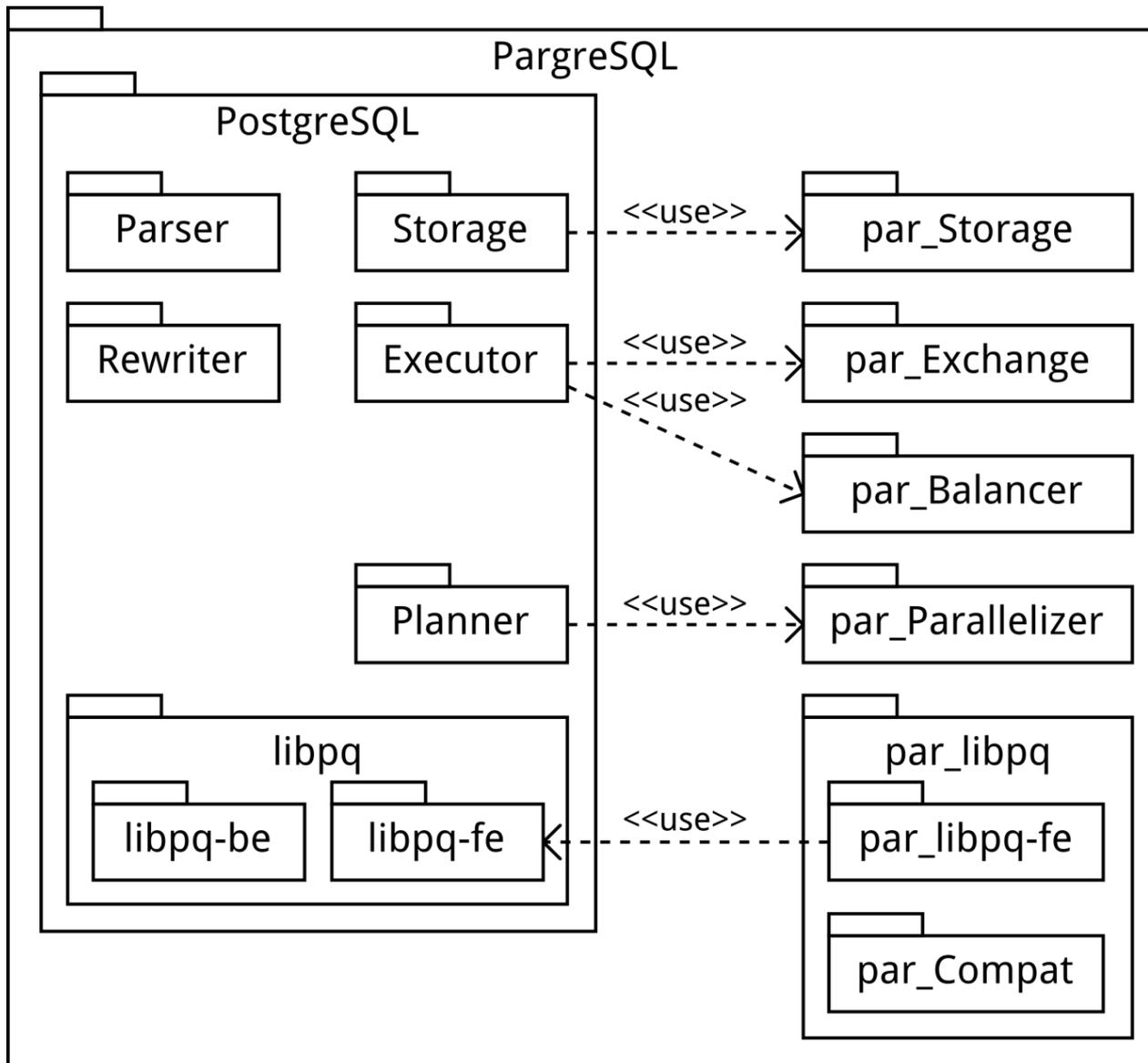
Оператор EXCHANGE



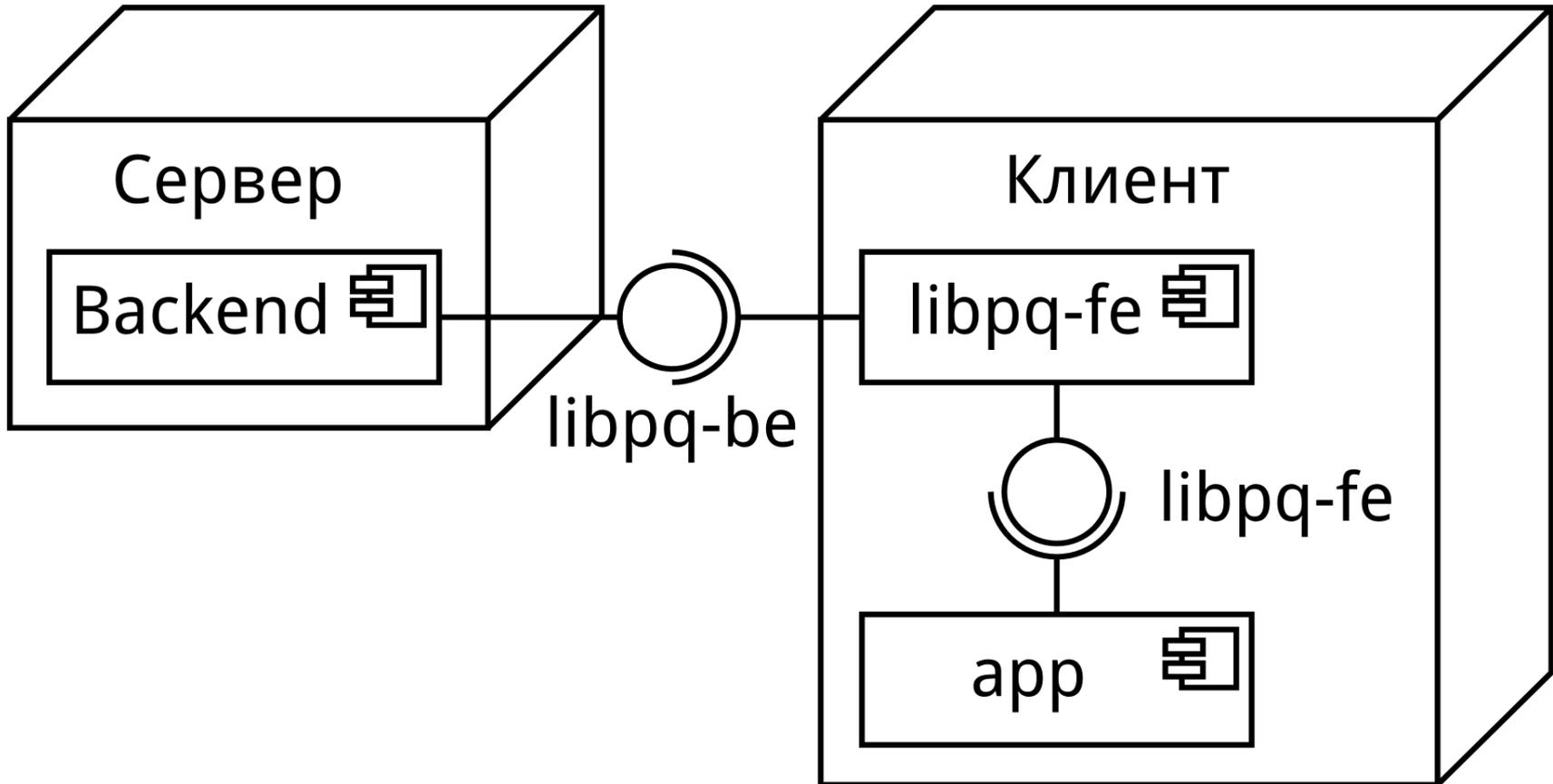
Модульная структура: PostgreSQL



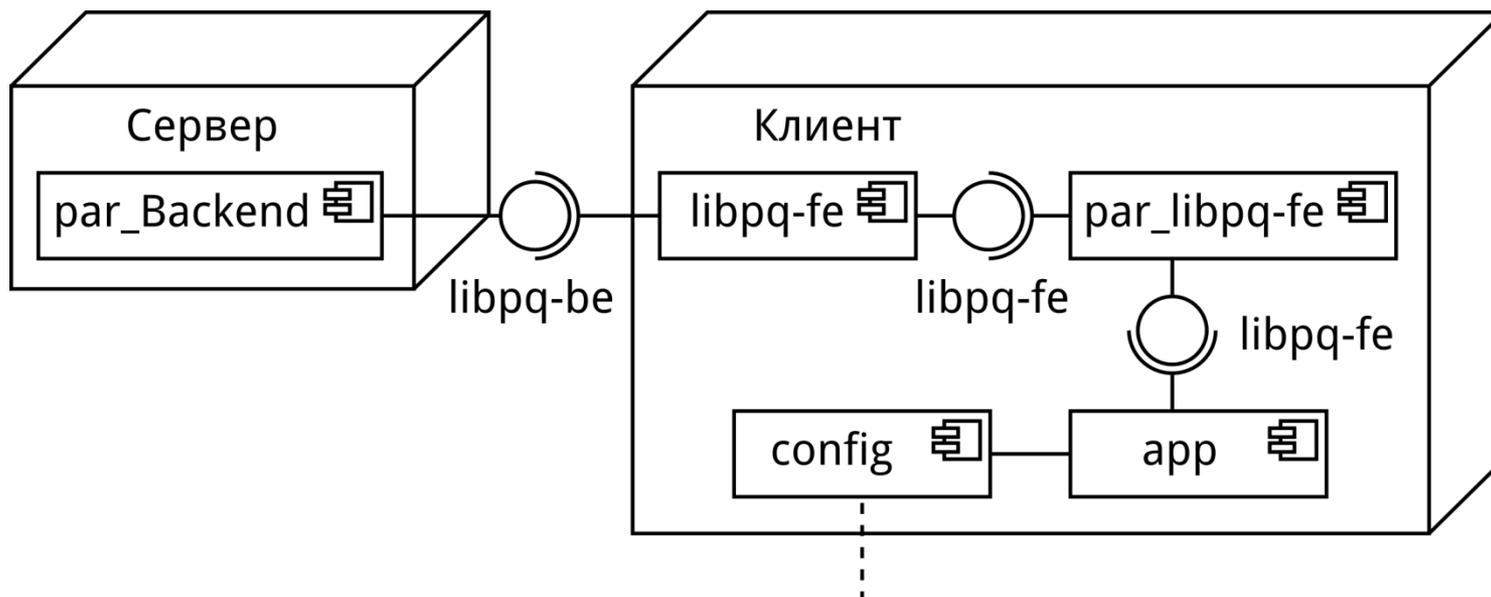
Модульная структура: PargreSQL



Размещение компонент: PostgreSQL



Размещение компонент: PargreSQL



```
<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>
```

Прозрачность перехода к PargreSQL

Приложение PostgreSQL

```
// app.c
#include <libpq-fe.h>

void main()
{
    PGconn c = PQconnectdb(...);
    PGresult r = PQexec(c, ...);
    ...
    PQfinish(c);
}
```



Приложение PargreSQL

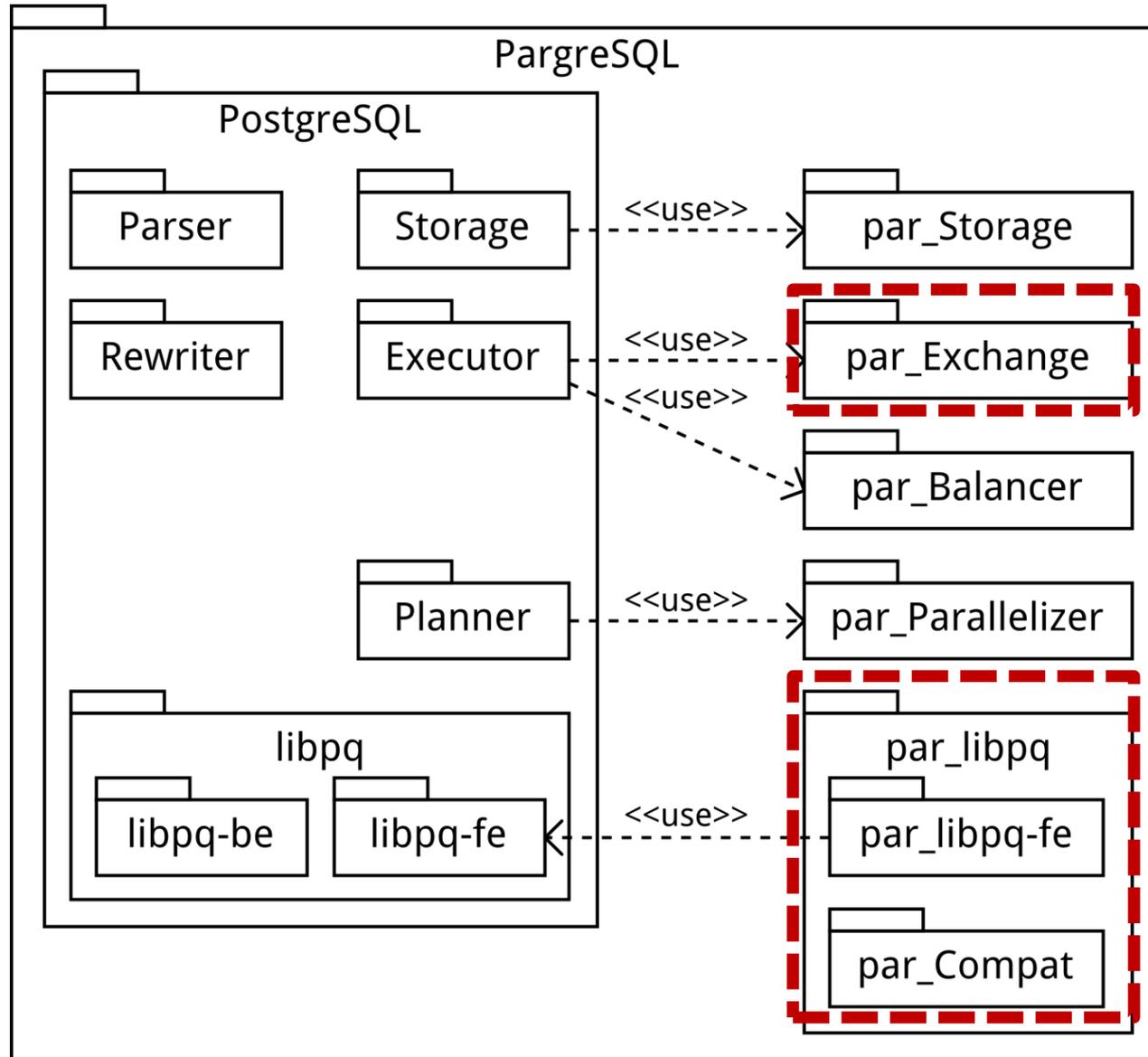
```
// par_app.c
#include <par_libpq-fe.h>

void main()
{
    PGconn c = PQconnectdb(...);
    PGresult r = PQexec(c, ...);
    ...
    PQfinish(c);
}
```

+

```
// par_Compact.h
#define PQconnectdb(...) \
    par_PQconnectdb(...)
#define PGconn \
    par_PGconn
...
```

Текущие результаты



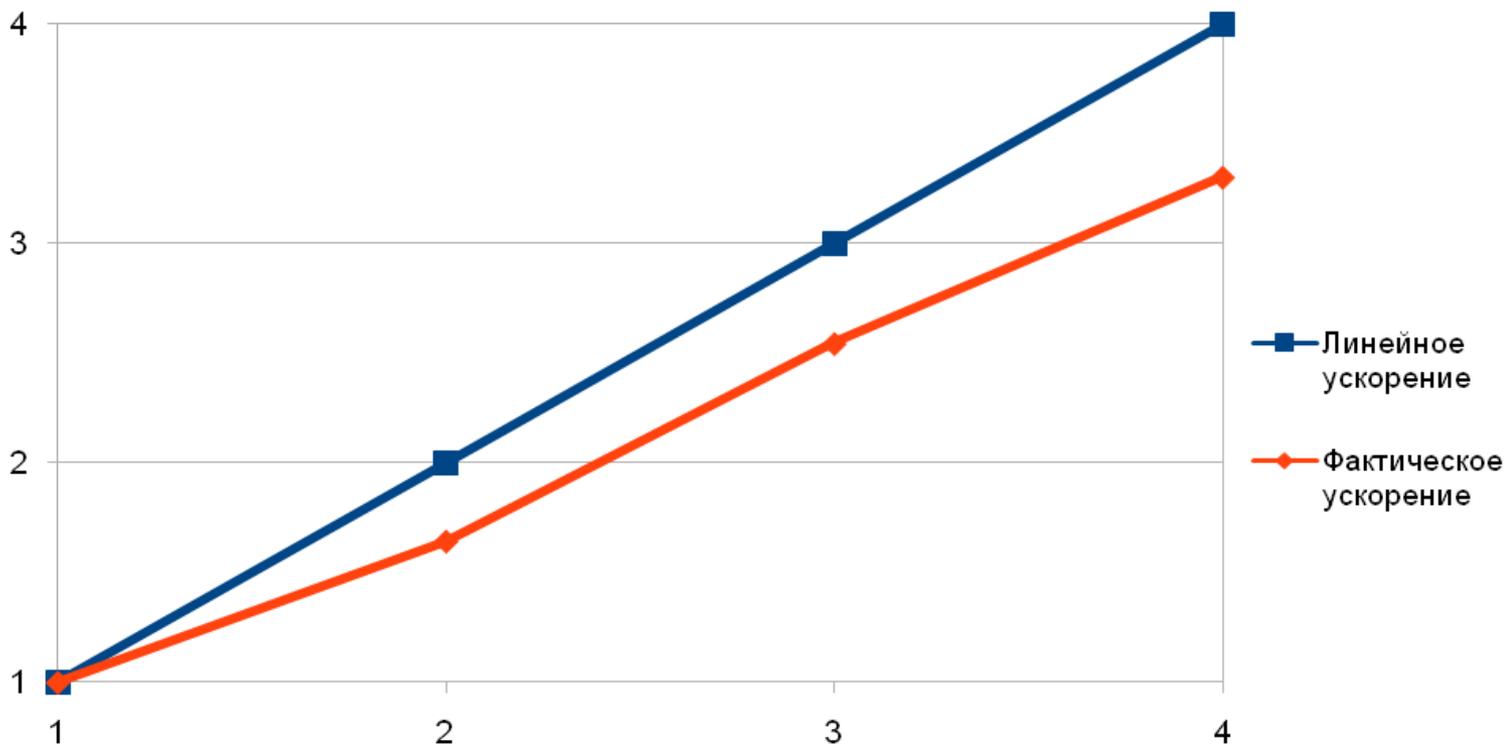
Текущие результаты

- Платформа: СКИФ Урал

- Лезвия/процессоры/ядра: 166/332/1328
- Процессор: Intel Xeon E5472
- Сеть: Infiniband (20 Гбит/с, 2 мкс)

- База данных

- 10^6 кортежей
- `select *`
`from T`
`where T.f % 10000 = 0`



Спасибо за внимание

- Вопросы?
 - Константин Сергеевич Пан
kpan@mail.ru
 - Михаил Леонидович Цымблер
mzym@susu.ru