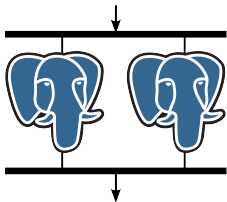


Второй Московский Суперкомпьютерный Форум
Москва, 27 октября 2011



Интеграция
параллелизма
в свободную СУБД
PostgreSQL

К.С. Пан, М.Л. Цымблер

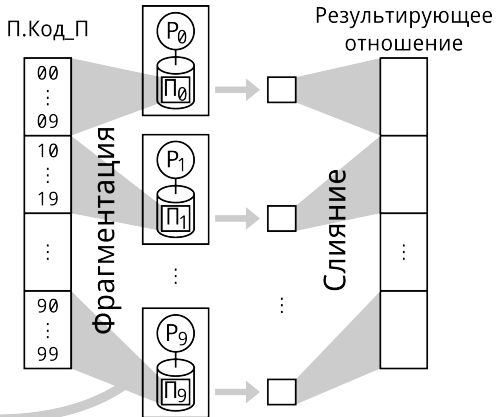
Южно-Уральский государственный университет
(национальный исследовательский университет), Челябинск

Работа выполнена при финансовой поддержке РФФИ (проект 09-07-00241-а)
и Минобрнауки РФ (госконтракт № 07.514.11.4036).

Фрагментный параллелизм

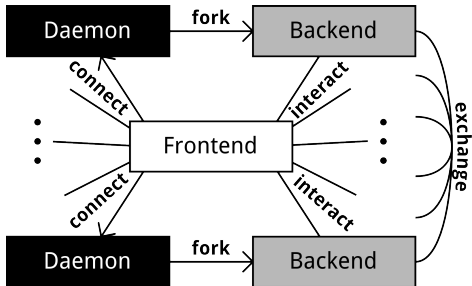
$$\Pi_i = \{t | t \in \Pi, \phi(t) = i\}$$
$$i = 0, \dots, 9$$

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

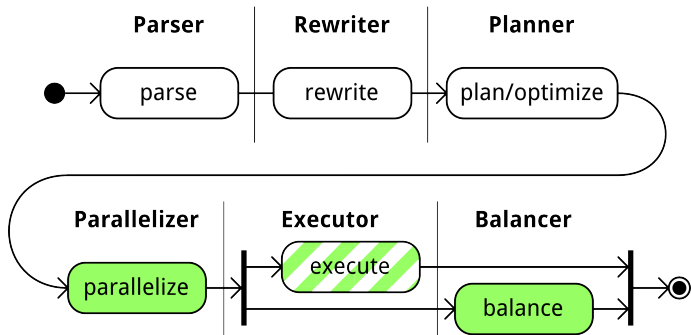


Архитектура

Процессы СУБД



Обработка запроса



Зачем нужны обмены

A	
1	40
2	30

B	
1	30
2	40

$C = A \text{ join } B \text{ on } A.2 = B.2$

C		
1	40	2
2	30	1

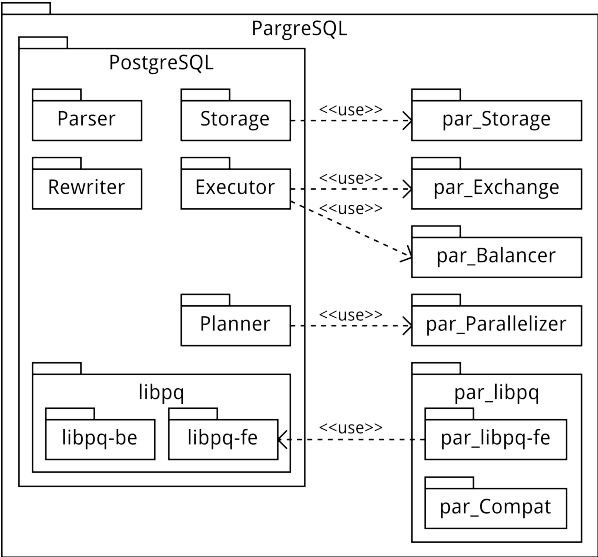
Зачем нужны обмены

A		B	
1	40	1	30
2	30	2	40

$C = A \text{ join } B \text{ on } A.2 = B.2$

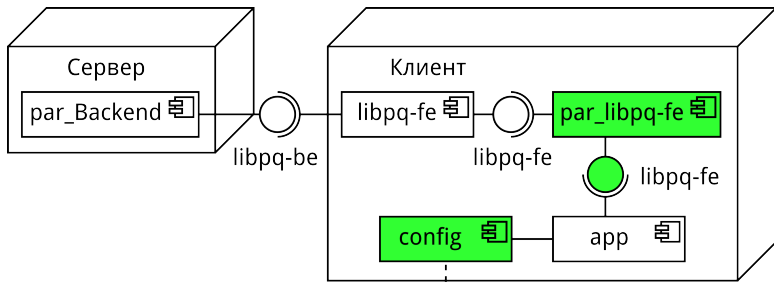
C	

Подсистемы



Реализация

Прикладная библиотека



```
<postgresql_config>  
  <node addr="10.0.0.40" port="5432"/>  
  <node addr="10.0.0.41" port="5432"/>  
</postgresql_config>
```

Прикладная библиотека

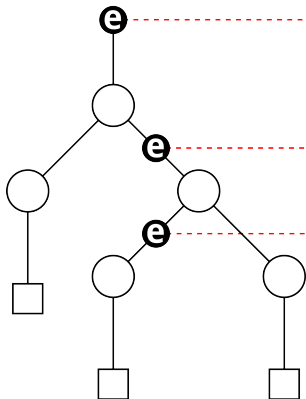
libpq-fe	par_libpq-fe
<pre>struct PGconn { ... };</pre>	<pre>struct par_PGconn { int len; struct PGconn *conns; };</pre>
<pre>PGconn *PQconnectdb(const char *conninfo);</pre>	<pre>par_PGconn *par_PQconnectdb() { for (...) { PQconnectdb(...); } }</pre>
<p>...</p>	<p>...</p>

Прикладная библиотека

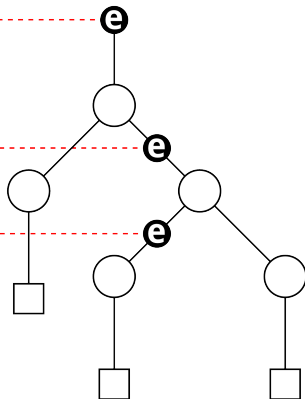
```
#define PGconn par_PGconn
#define PQconnectdb(X) par_PQconnectdb()
#define PQfinish(X) par_PQfinish(X)
#define PQstatus(X) par_PQstatus(X)
#define PQexec(X,Y) par_PQexec(X,Y)
```

Оператор Exchange

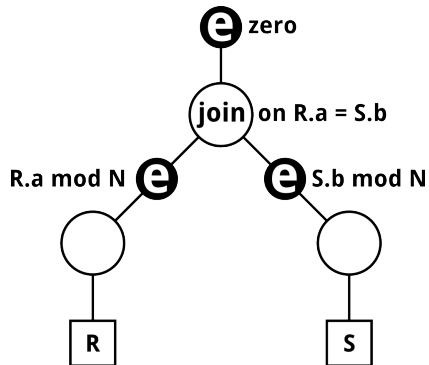
Выч. узел 0



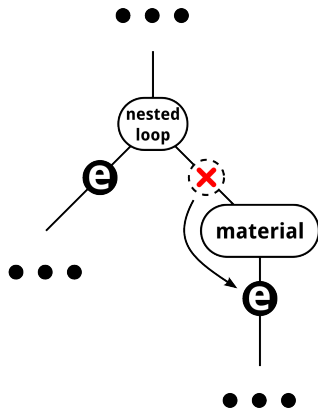
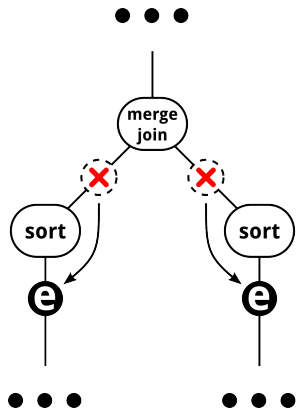
Выч. узел 1



Параллелизатор



Параллелизатор



Фрагментация таблиц

`pg_class.reloptions` — атрибуты хранения отношения.

Синтаксис управления `reloptions`:

-- создать таблицу с указанными атрибутами фрагментации

```
CREATE TABLE t1 (...) WITH (fragattr = 3, ...);
```

-- изменить атрибуты фрагментации

```
ALTER TABLE t1 SET (fragattr = 3);
```

```
ALTER TABLE t1 RESET (fragattr);
```


Вычислительные эксперименты



Платформа:

Кластер СКИФ Урал
2 x Intel Xeon E5472 на узел
сеть InfiniBand 20 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
```

Продвижение проекта

Запланировано включение PargreSQL в официальную версию PostgreSQL.

Спасибо

Константин Пан
kvaren@gmail.com

Михаил Цымблер
zymbler@gmail.com