

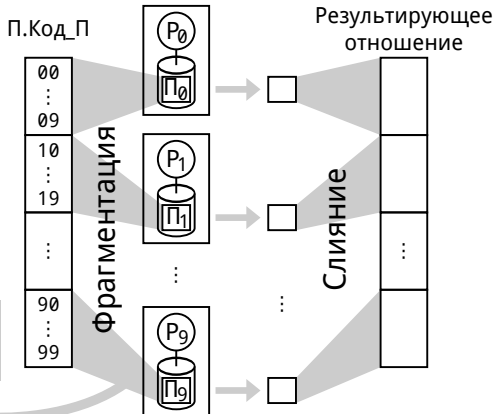
Разработка
параллельной СУБД
на основе свободной
СУБД PostgreSQL

К.С. Пан
20 сентября 2011

Фрагментный параллелизм

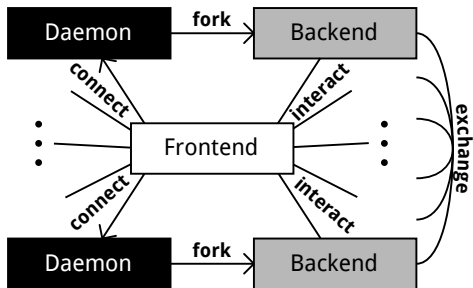
$$\Pi_i = \{t | t \in \Pi, \phi(t) = i\}$$
$$i = 0, \dots, 9$$

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

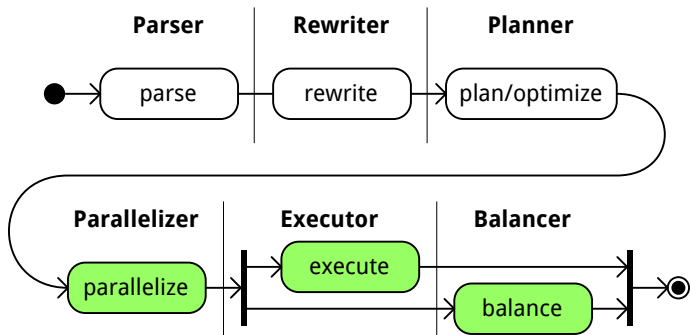


Архитектура

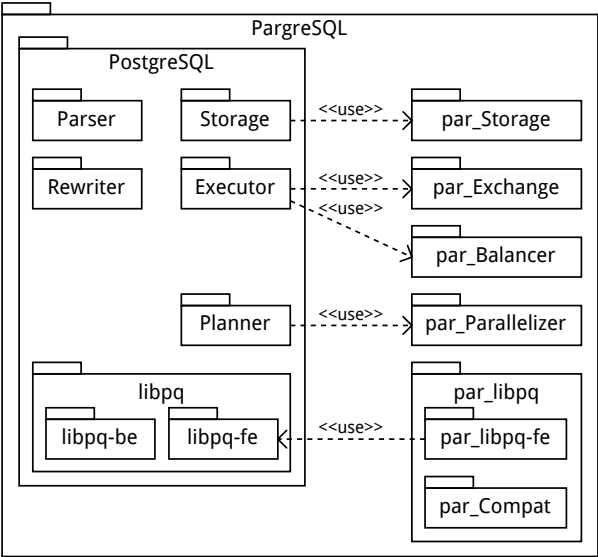
Процессы в СБД



Обработка запроса

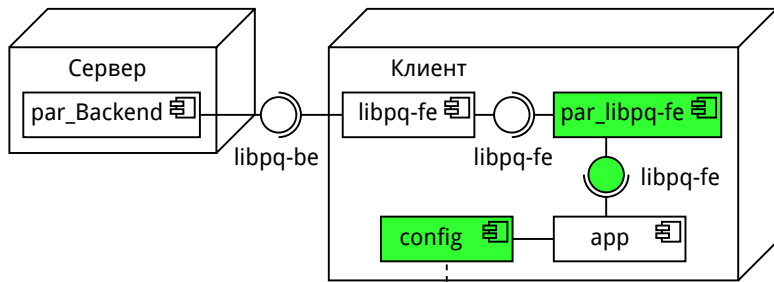


Подсистемы



Реализация

Прикладная библиотека



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


Прикладная библиотека

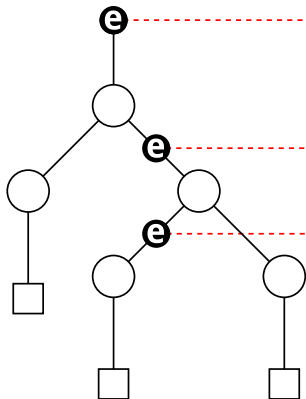
libpq-fe	par_libpq-fe
<pre>struct PGconn { ... };</pre>	<pre>struct par_PGconn { int len; struct PGconn *conns; };</pre>
<pre>PGconn *PQconnectdb(const char *conninfo);</pre>	<pre>par_PGconn *par_PQconnectdb() { for (...) { PQconnectdb(...); } }</pre>
...	...

Прикладная библиотека

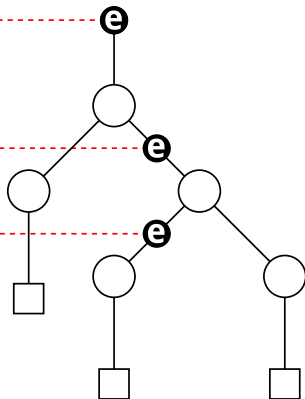
```
#define PGconn par_PGconn
#define PQconnectdb(X) par_PQconnectdb()
#define PQfinish(X) par_PQfinish(X)
#define PQstatus(X) par_PQstatus(X)
#define PQexec(X,Y) par_PQexec(X,Y)
```

Оператор Exchange

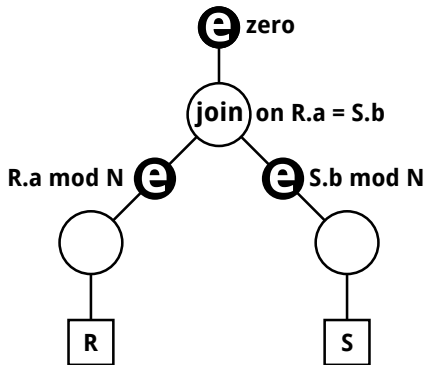
Выч. узел 0



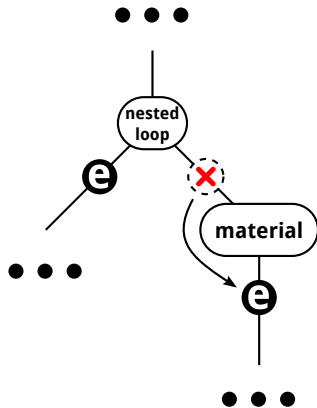
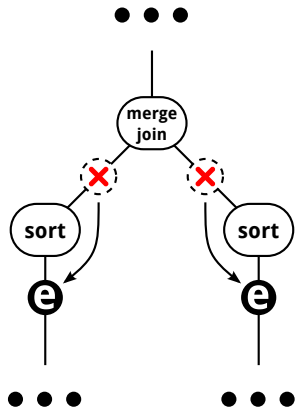
Выч. узел 1



Параллелизатор



Параллелизатор



Метаданные

`pg_class.reloptions` — атрибуты хранения отношения.

Синтаксис управления `reloptions`:

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

Спасибо

Константин Пан
kvaren@gmail.com