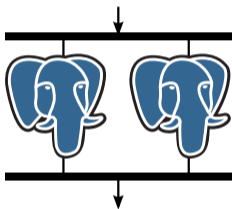


PGCONF.RUSSIA 2015

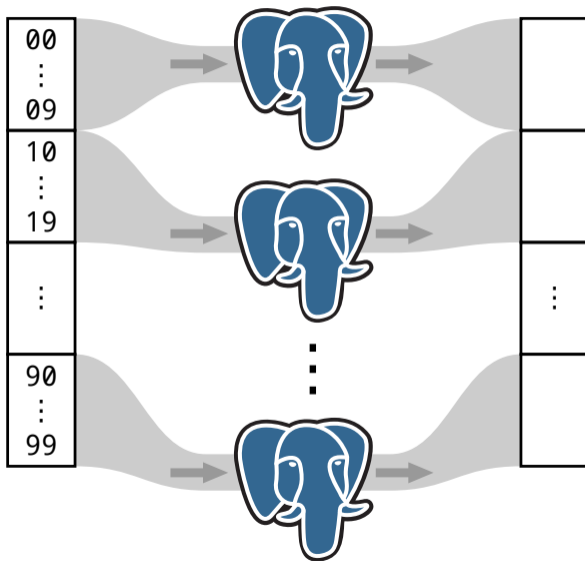
Москва, 6 февраля



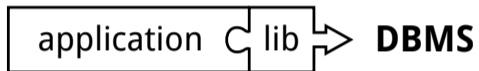
Ещё один  
параллельный  
PostgreSQL

Константин Пан  
ЮУрГУ, Челябинск

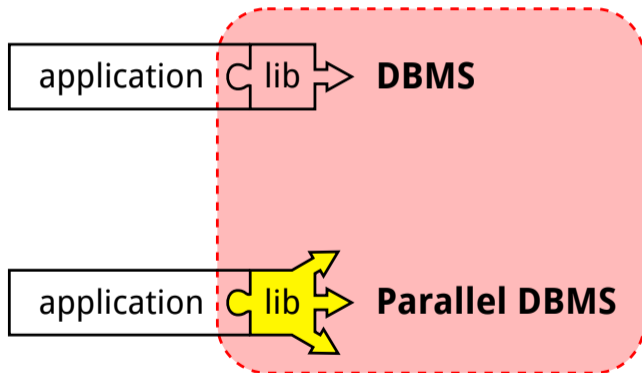
## Горизонтальная фрагментация



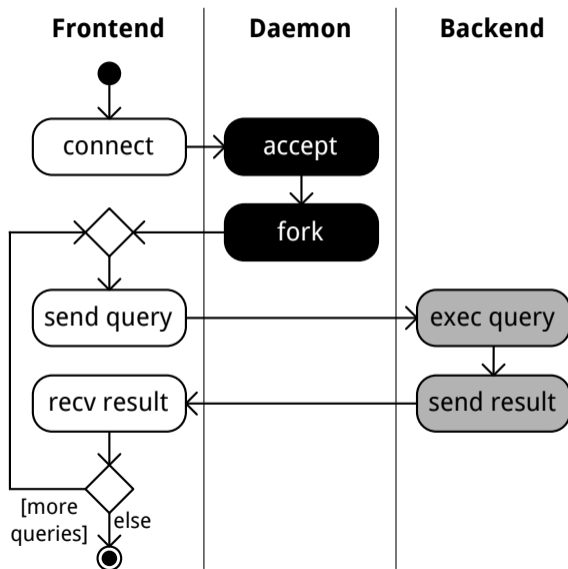
## Приложение параллельной БД



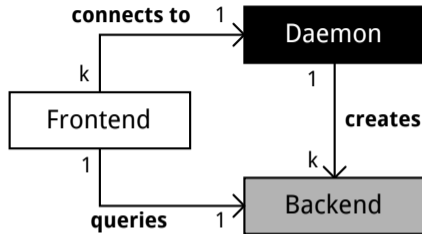
## Приложение параллельной БД



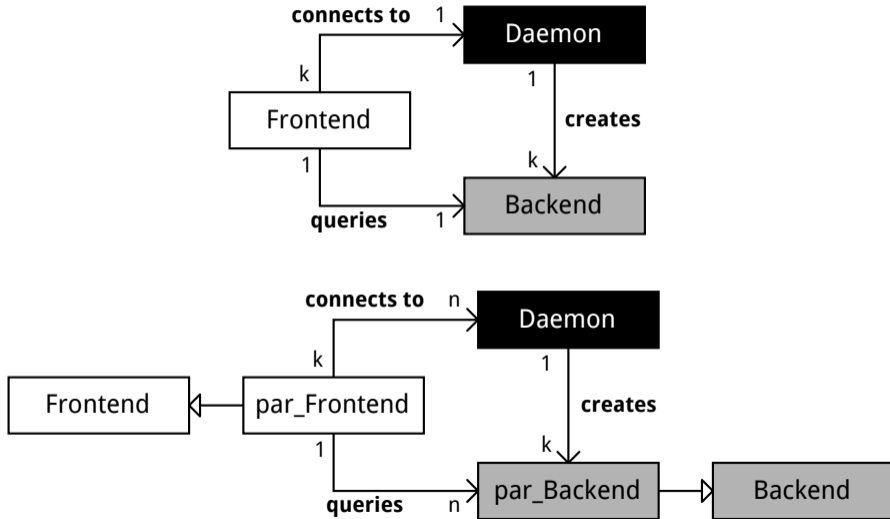
## Взаимодействие клиента и сервера



## Взаимодействие клиента и сервера



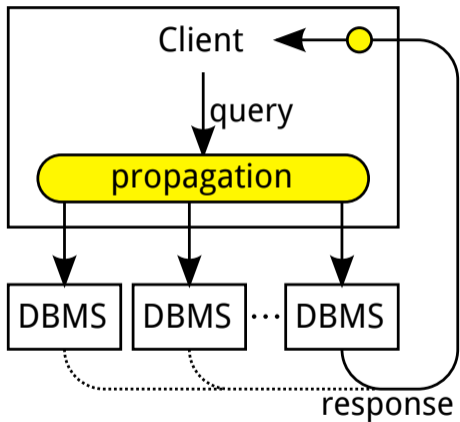
## Взаимодействие клиента и сервера



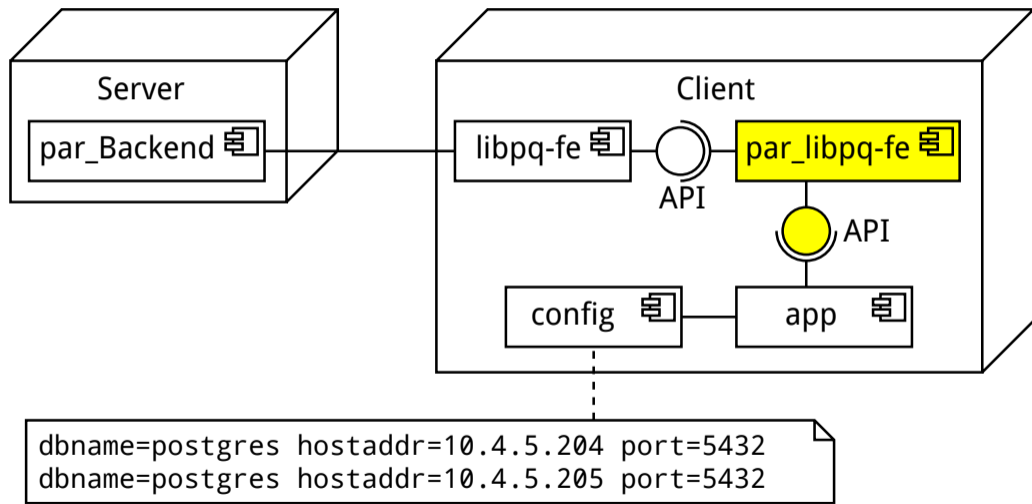
par\_Frontend



# Прикладная библиотека



## Реализация прикладной библиотеки

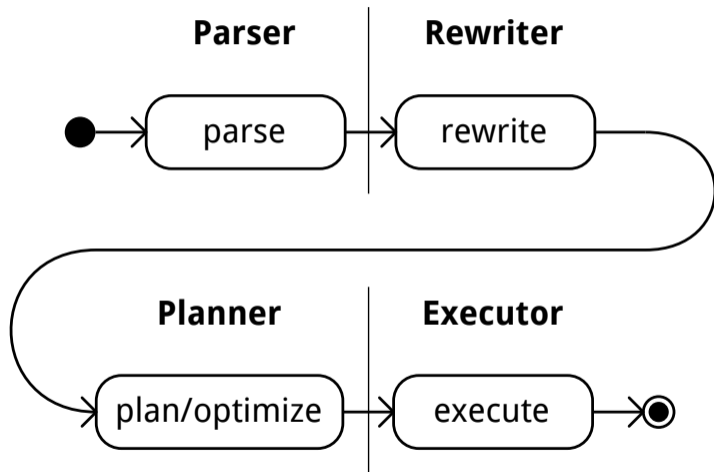


## Реализация прикладной библиотеки

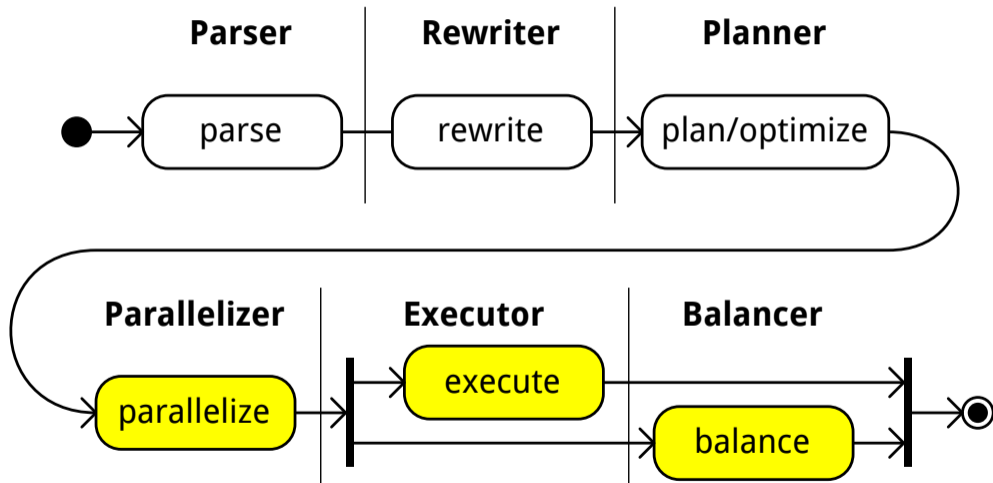
```
//      original          parallel
#define PGconn            par_PGconn
#define PQconnectdb(X)   par_PQconnectdb()
#define PQfinish(X)      par_PQfinish(X)
#define PQstatus(X)      par_PQstatus(X)
#define PQexec(X,Y)      par_PQexec(X,Y)
```

par\_Backend

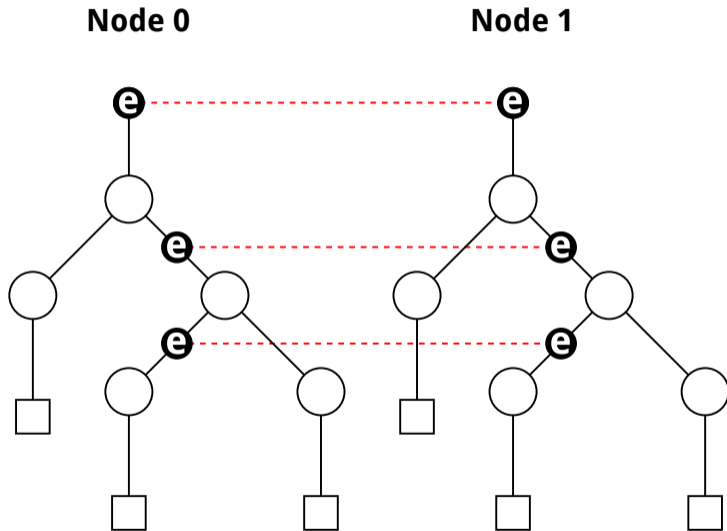
## Исполнение запросов



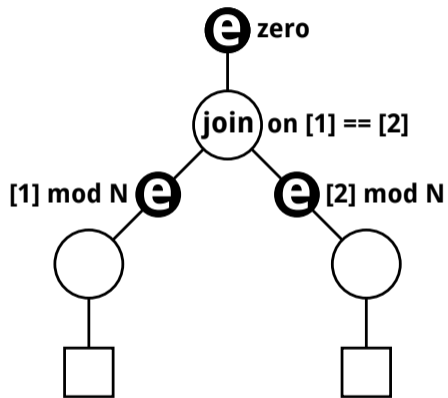
## Параллельное исполнение запросов



## Параллельный план запроса

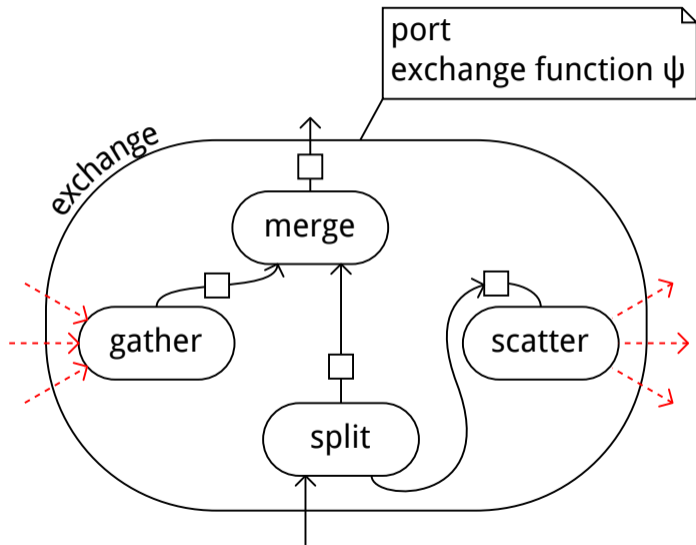


# Параллелизация

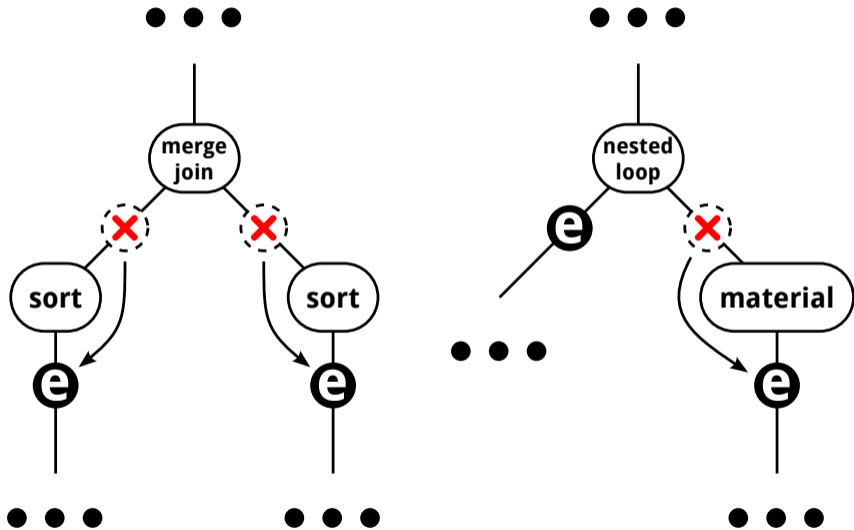




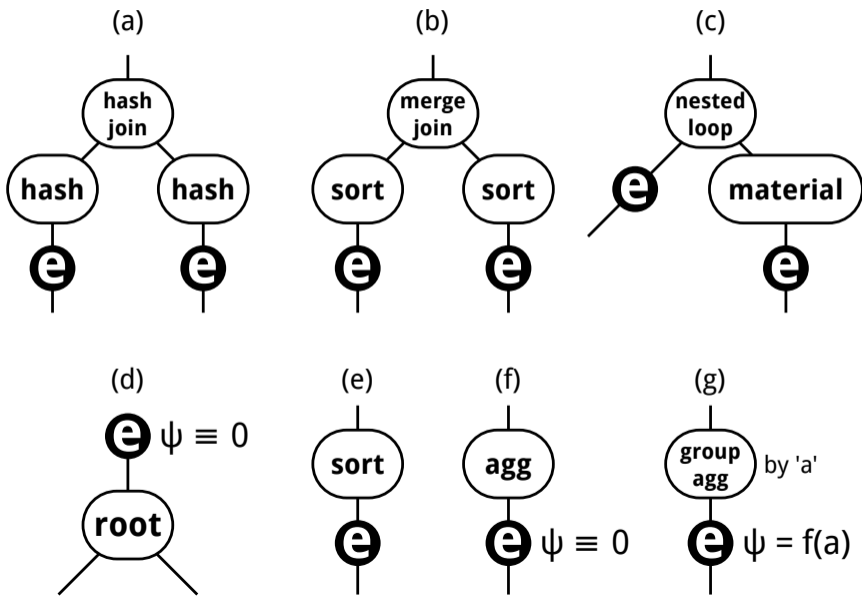
## Операция обмена



## Обмены при Join в PostgreSQL



# Куда вставляется Exchange?

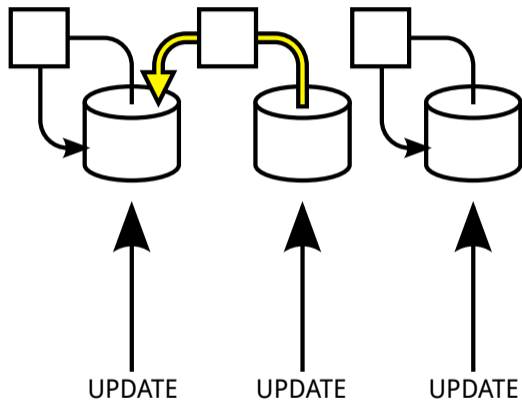
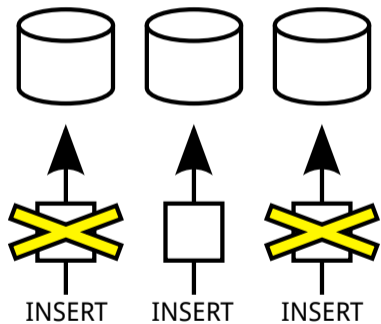


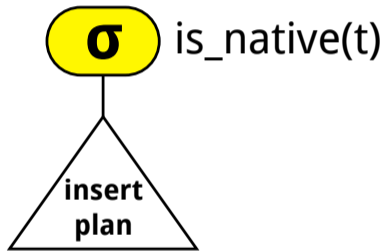
## Метаданные о фрагментации

- ▶ указываются в ходе создания таблицы
- ▶ задают атрибут  $x$  для функции фрагментации  $\varphi(t) = t.x \% N$

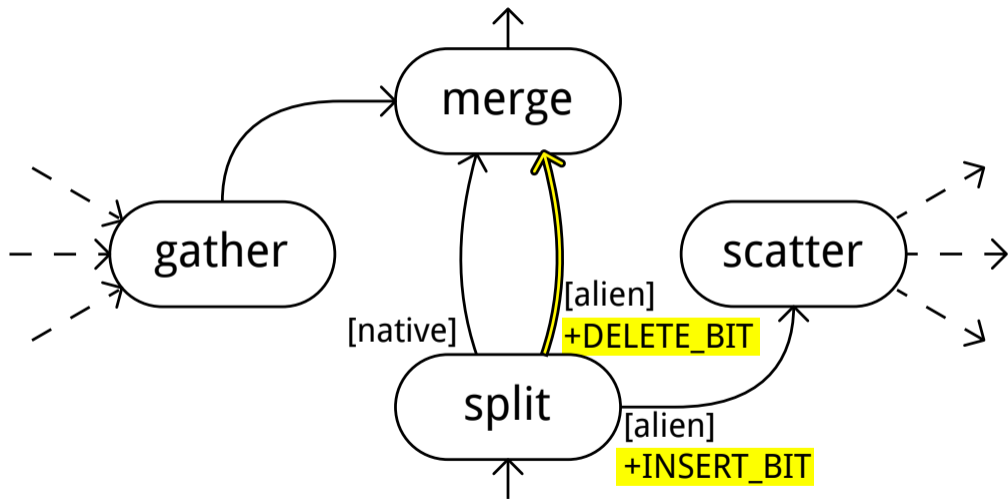
```
create table Person (  
    id int,  
    name varchar(30),  
    gender char(1),  
    birth date  
) with (fragattr = id);
```

## Запросы на изменение данных



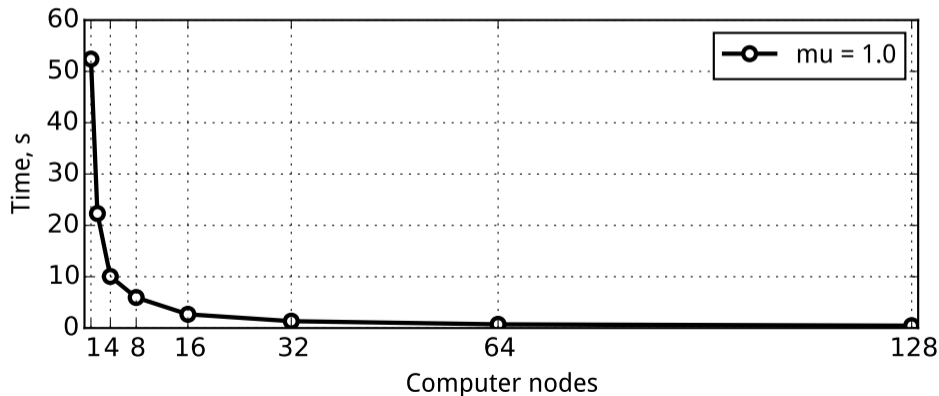


# Update



## Эксперименты

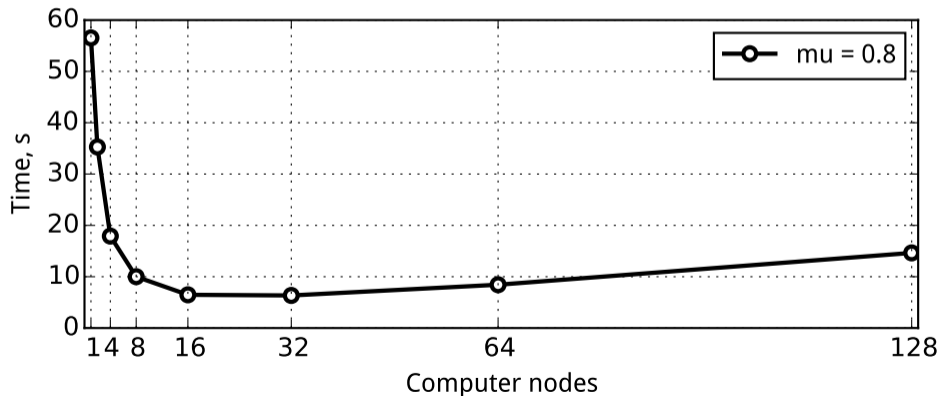
- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении





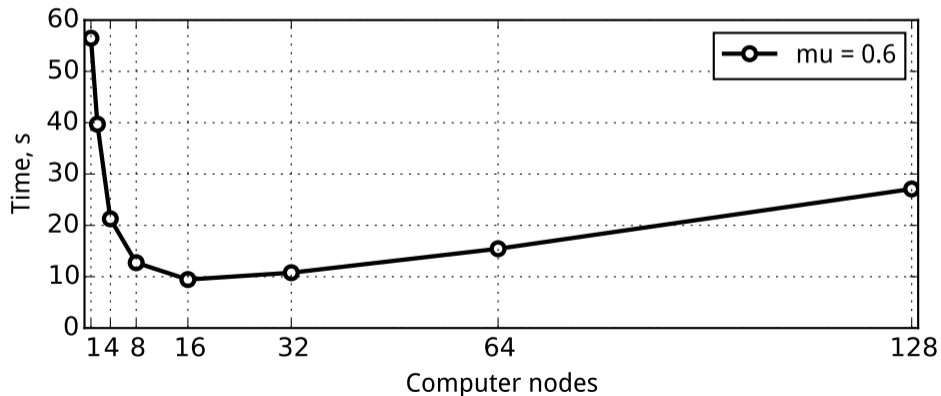
## Эксперименты

- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении



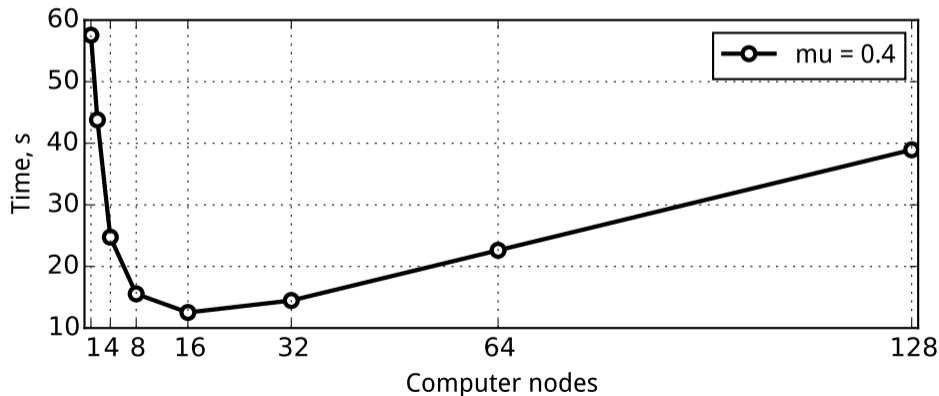
## Эксперименты

- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении



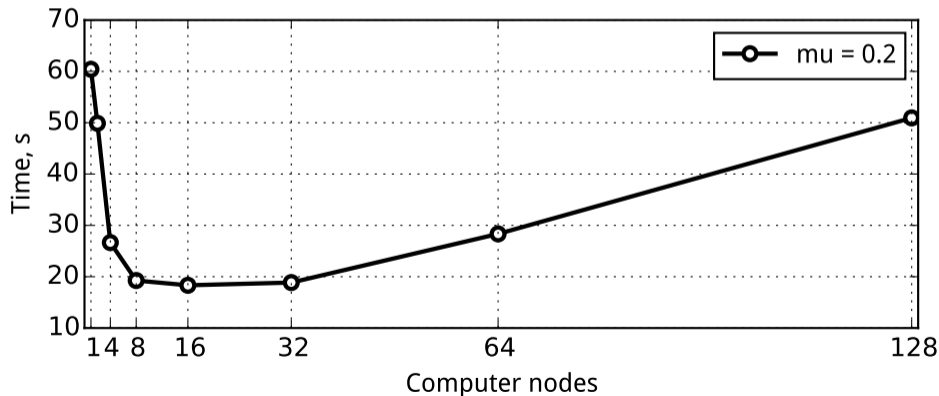
## Эксперименты

- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении



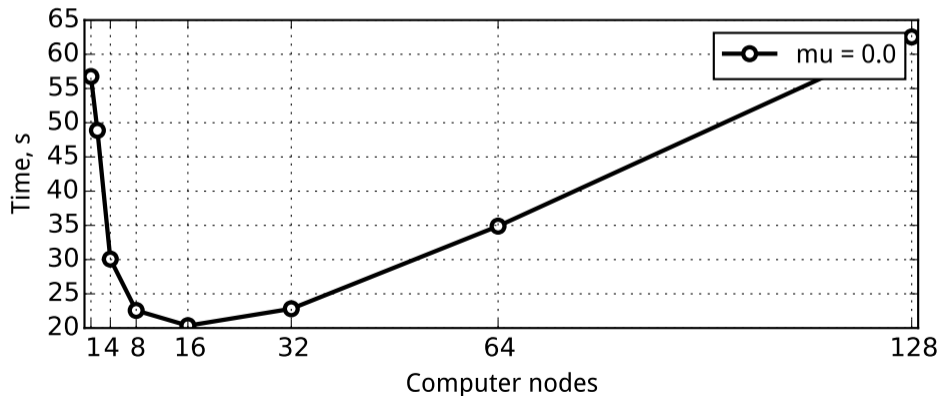
## Эксперименты

- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении



## Эксперименты

- ▶ `/select * from R, S where R.c = S.c and R.a % 100003=0`
- ▶  $|R| = 60 \cdot 10^6$ , fragattr = b
- ▶  $|S| = 1.5 \cdot 10^6$ , fragattr = b
- ▶  $\mu$  — количество «своих» кортежей при соединении



## Нерешённые вопросы

- ▶ Согласование планов между узлами
- ▶ Плохая реализация передачи данных
- ▶ Подзапросы и триггеры
- ▶ Индексы
- ▶ Использование не только числовых атрибутов для фрагментации
- ▶ Использование произвольно заданной функции для фрагментации
- ▶ Агрегация без группировки

Спасибо за внимание

Constantin S. Pan  
kvapen@gmail.com