

*Международная Суперкомпьютерная конференция  
"Научный сервис в сети Интернет: все грани параллелизма"  
23-28 сентября 2013 г., Новороссийск*

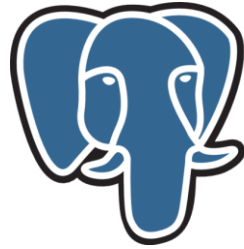
# Параллельная СУБД PargreSQL: исследование эффективности

М.Л. Цымблер, К.С. Пан  
Южно-Уральский государственный университет  
(Челябинск)



Работа выполнена при финансовой поддержке РФФИ  
(гранты №№ 12-07-00443-а, 12-07-31217 мол\_а)

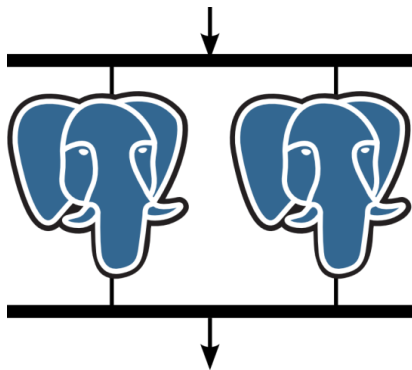
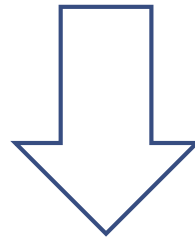
# СУБД PargreSQL



PostgreSQL

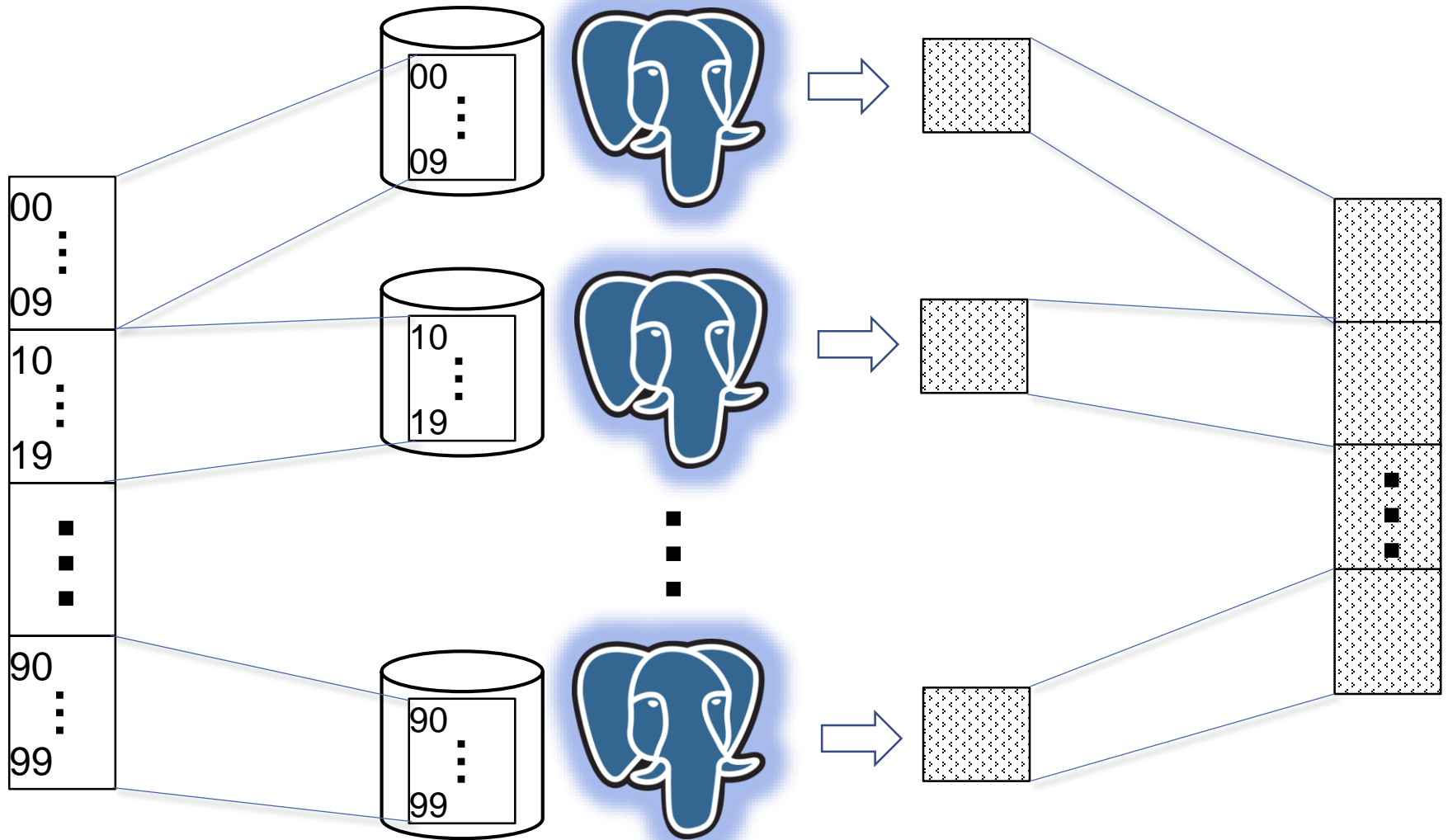
+

**фрагментный параллелизм**

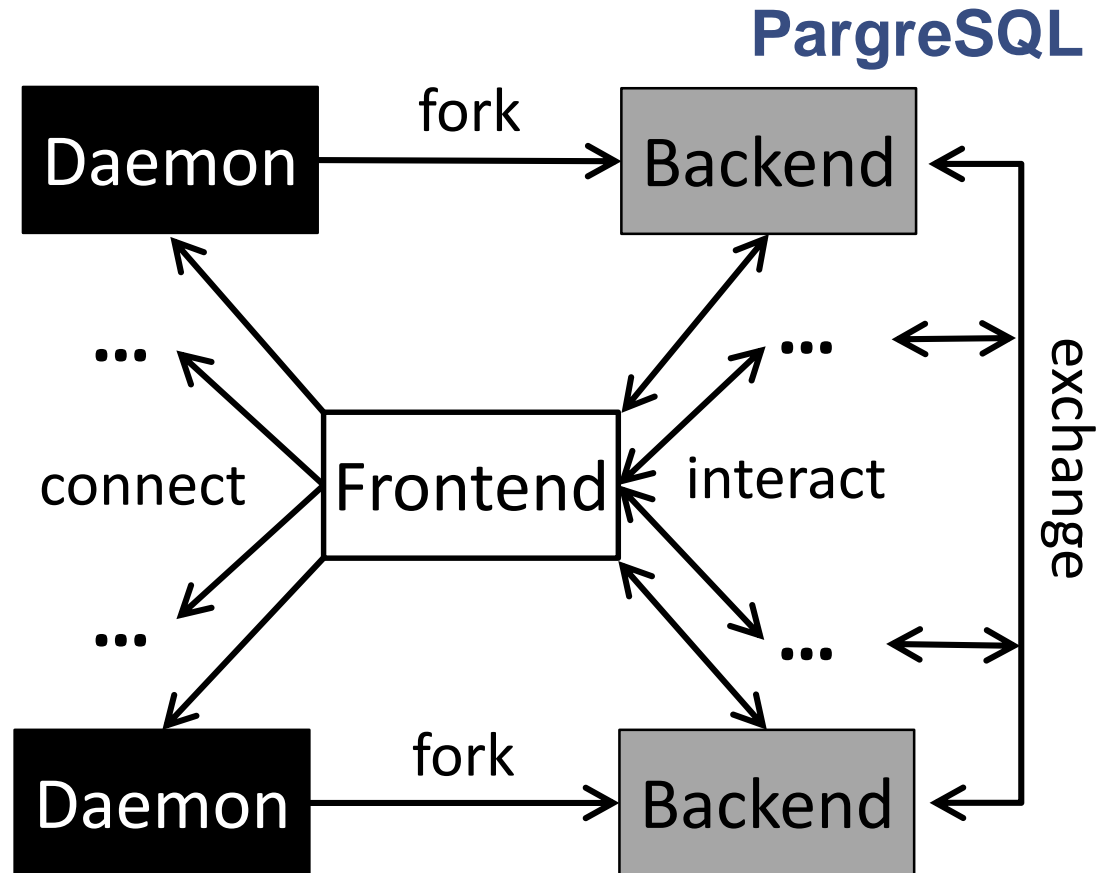
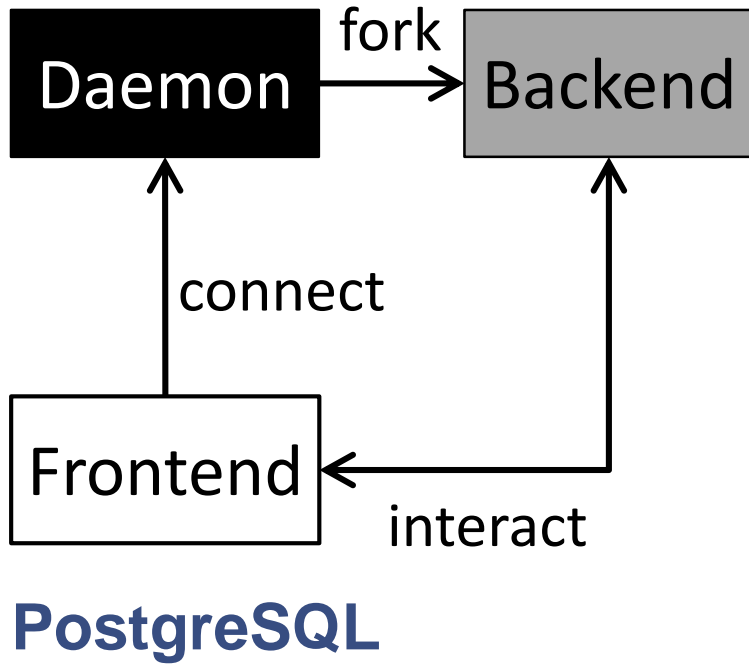


PargreSQL

# Фрагментный параллелизм

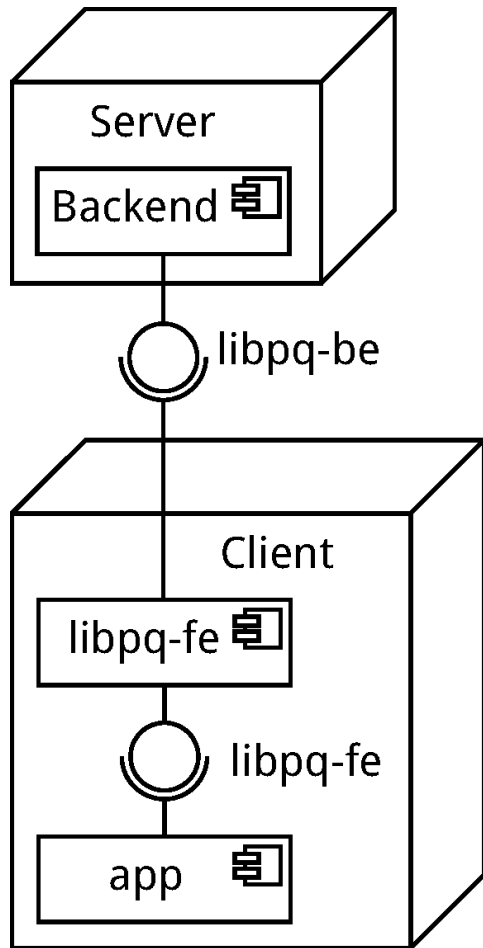


# Процессы СУБД

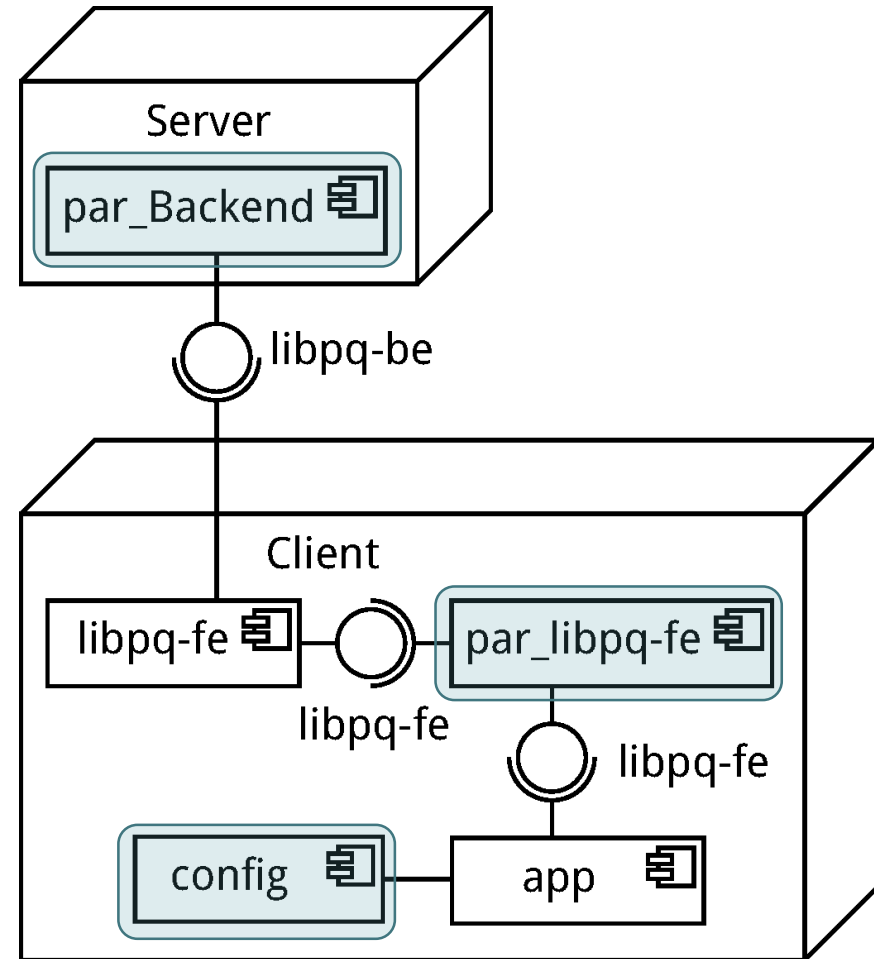


# Размещение компонентов

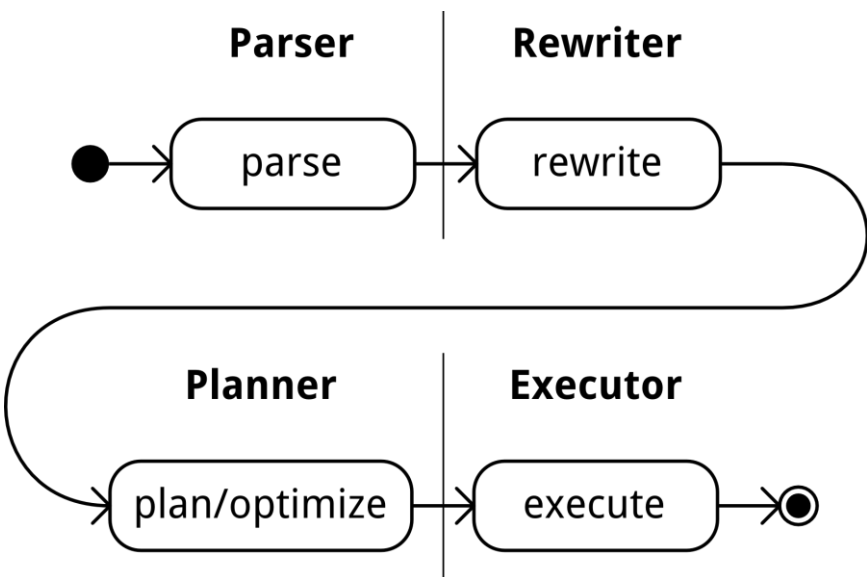
## PostgreSQL



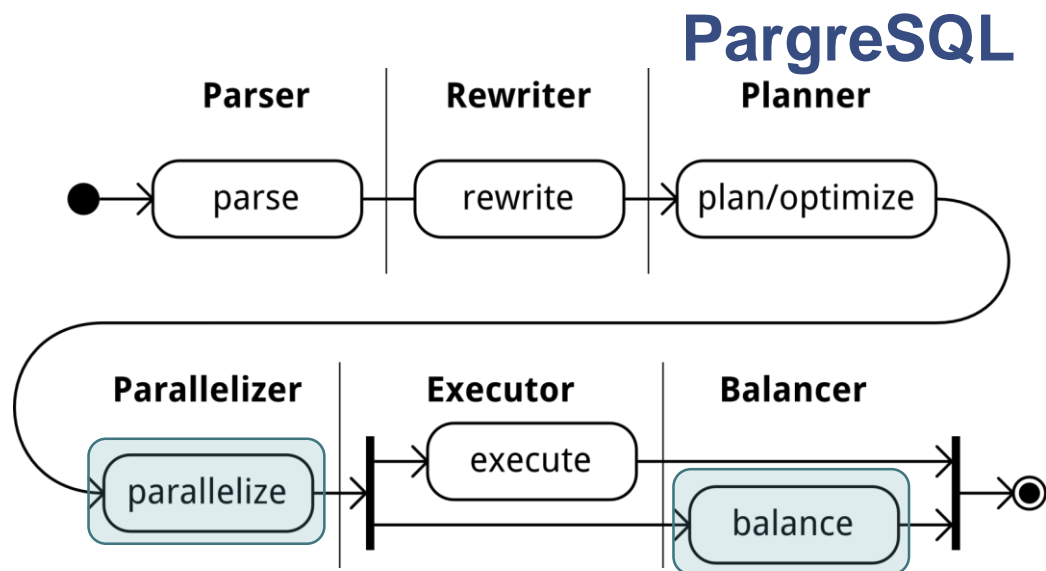
## PargreSQL



# Обработка запроса



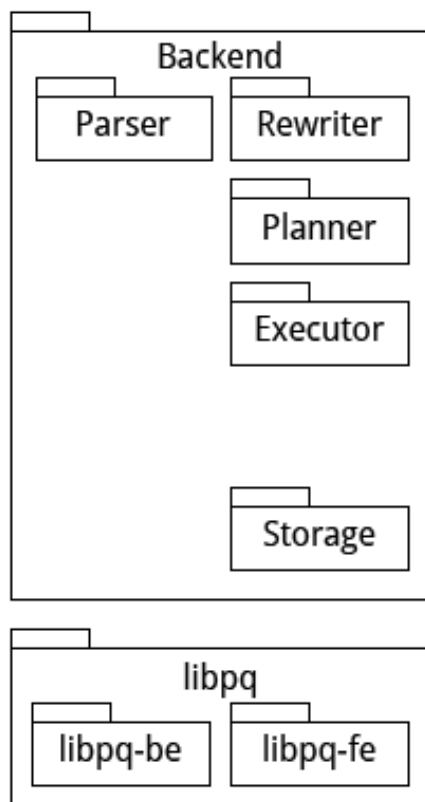
PostgreSQL



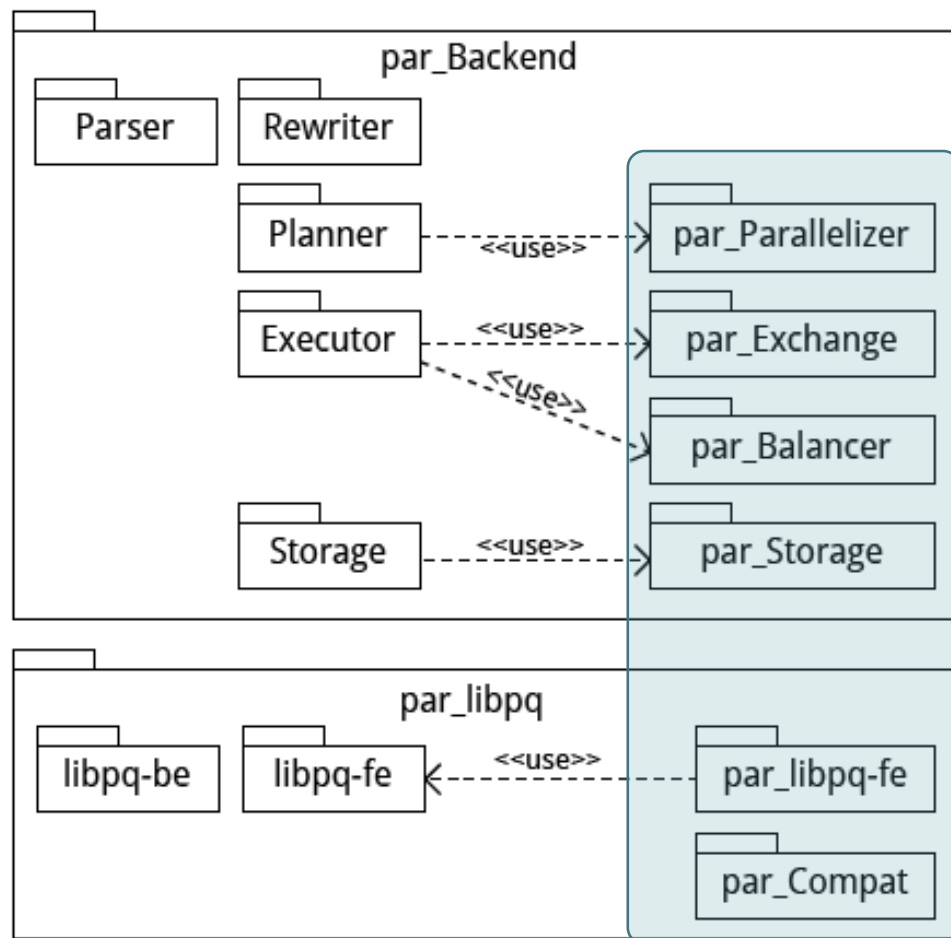
PargreSQL

# Архитектура СУБД

## PostgreSQL



## PargreSQL



# Фрагментация таблиц

```
create table Авторы (  
    АвторИД integer primary key,  
    ФИО ...  
    ...)
```

```
with (fragattr = АвторИД) ;
```

```
-- Атрибут фрагментации: АвторИД  
-- Функция фрагментации: АвторИД mod N  
-- (N – количество узлов кластера)
```



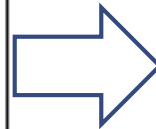
# Совместимость Postgres→Pargres

## par\_Compat

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

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

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



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

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

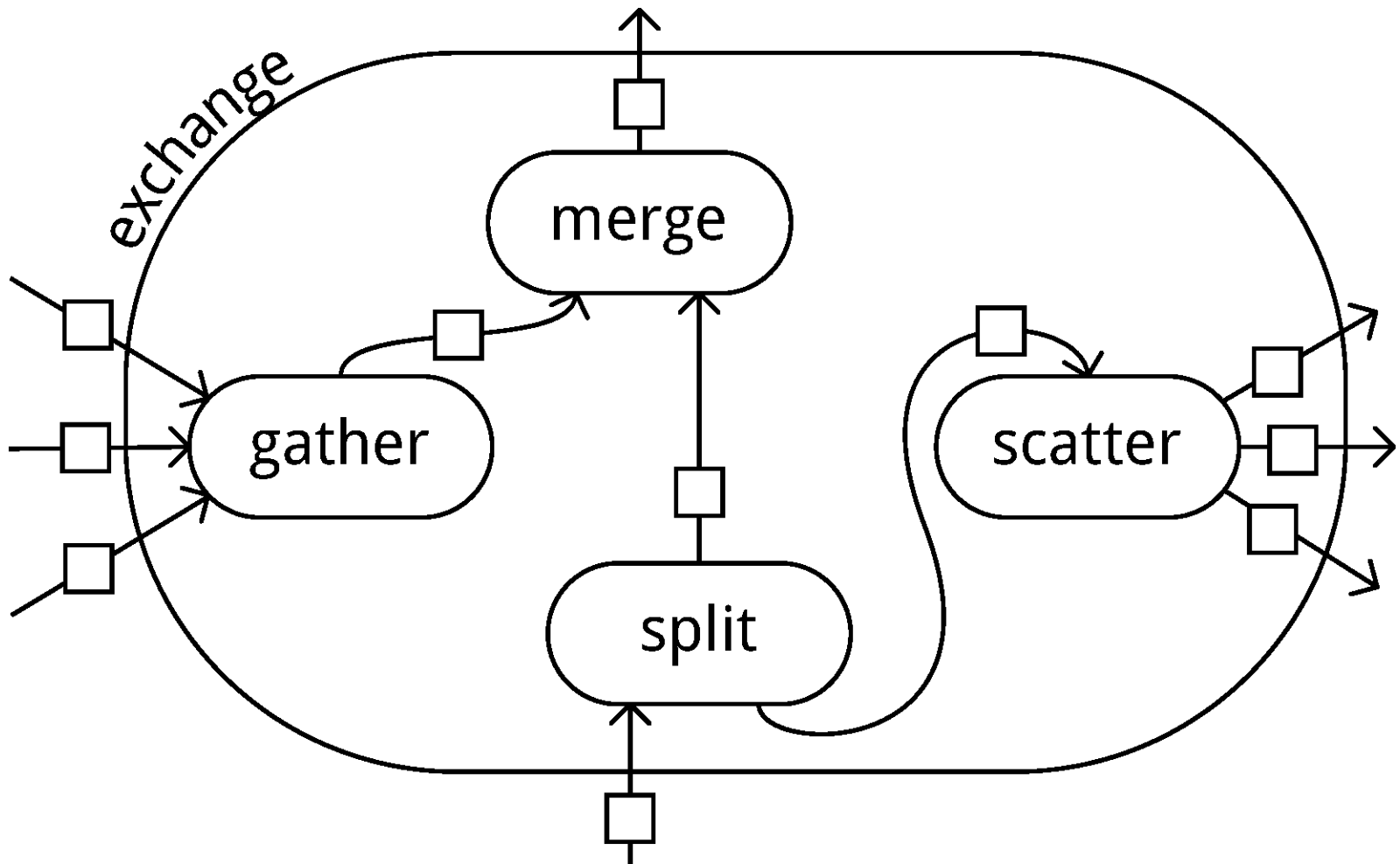
**PostgreSQL приложение**

Научный сервис в сети Интернет 2013

**PargresSQL приложение**

25.09.2013

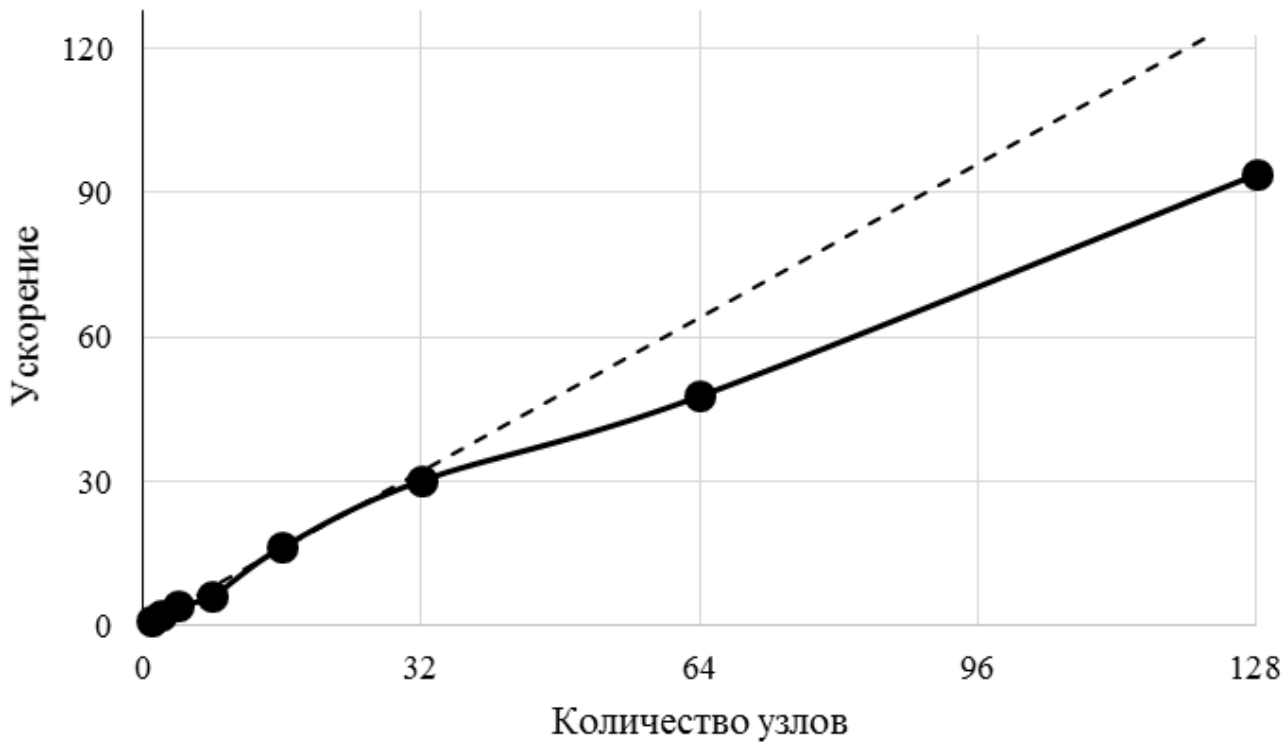
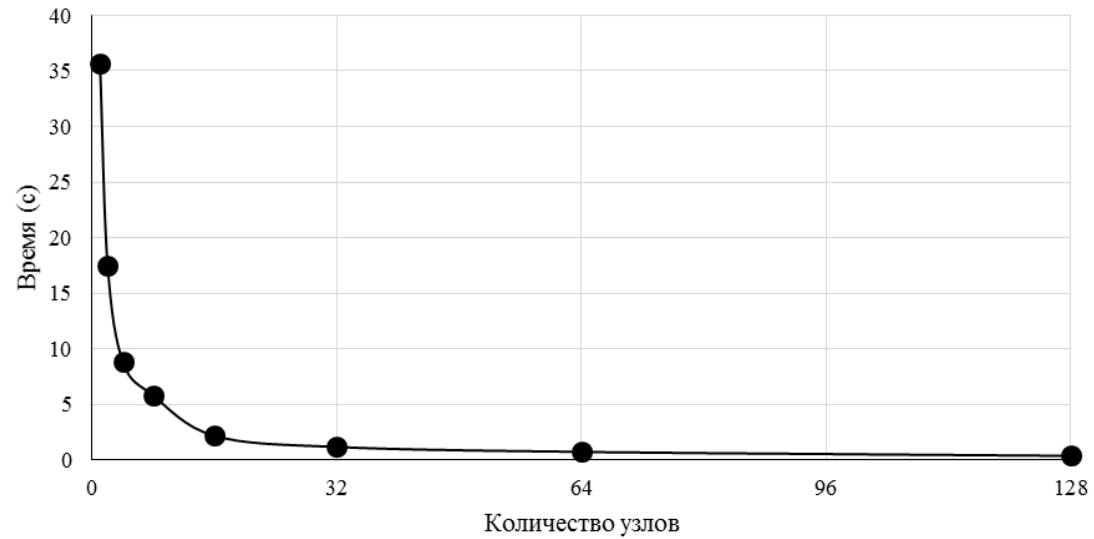
# Оператор EXCHANGE



# Эксперименты

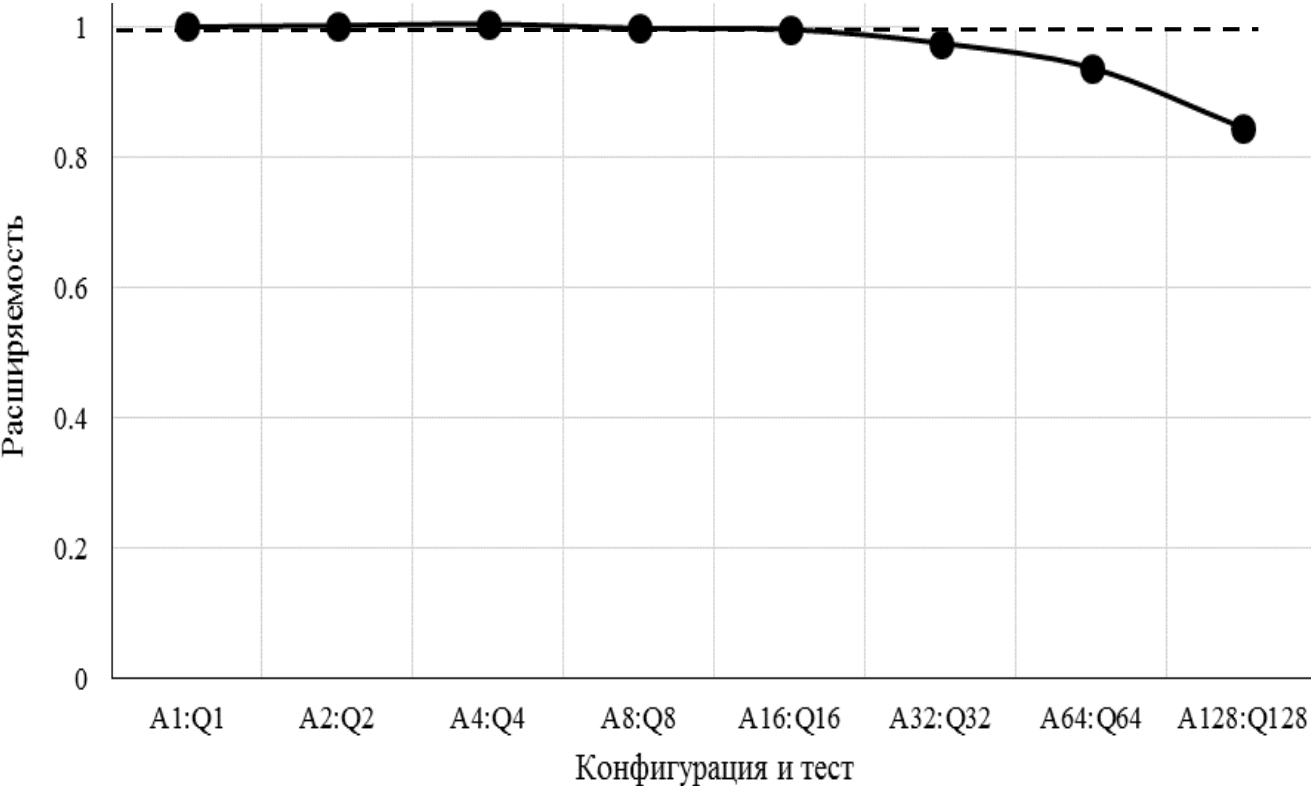
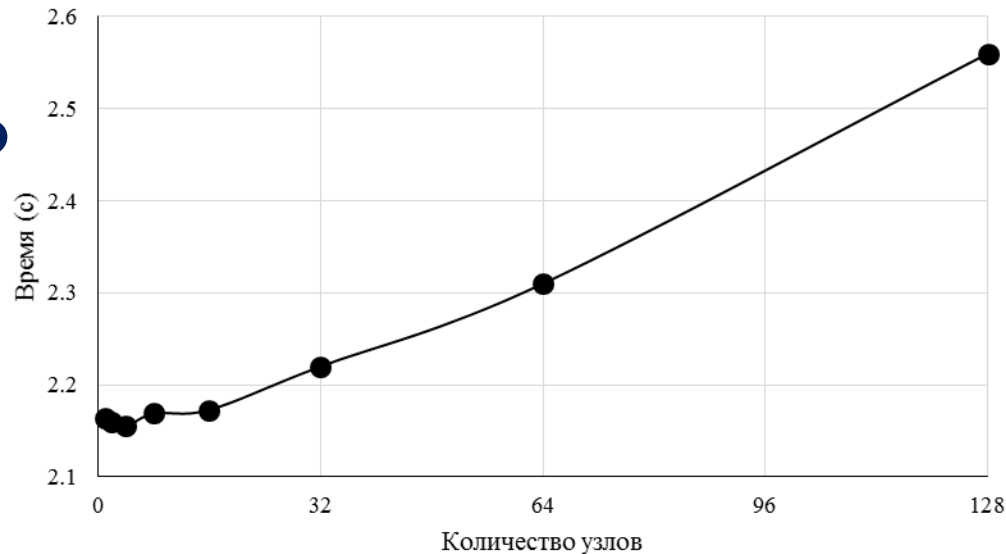
- Ускорение и расширяемость
  - `select * from R, S where R.c=S.c`
- Обработка транзакций
  - Тест ТРС-С: имитация работы склада
- Сверхбольшие данные
  - Разбиение графов
- Аппаратная платформа
  - Суперкомпьютер Tornado ЮУрГУ

# Ускорение



$$|R|=3 \cdot 10^8$$
$$|S|=7.5 \cdot 10^5$$

# Расширяемость



**A1:Q1**  
 $|R|=1.2 \cdot 10^6$   
 $|S|=3 \cdot 10^5$

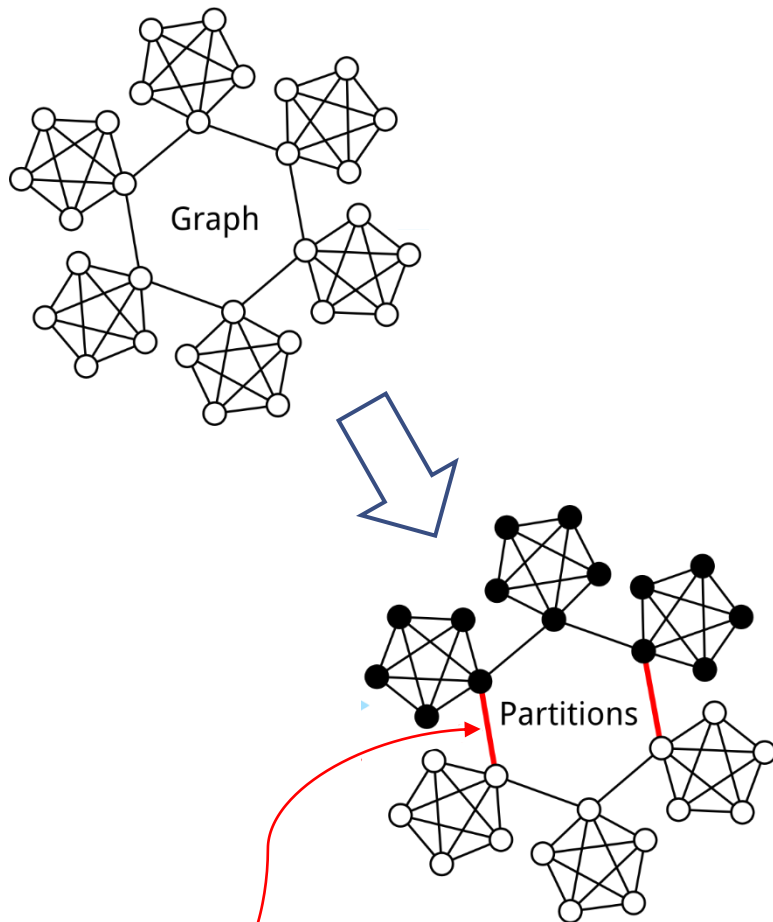
...

**A128:Q128**  
 $|R|=1.5 \cdot 10^8$   
 $|S|=3.8 \cdot 10^6$

# Эксперименты: тест TPC-C

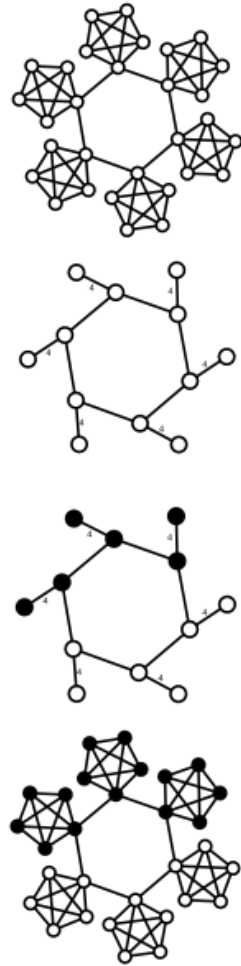
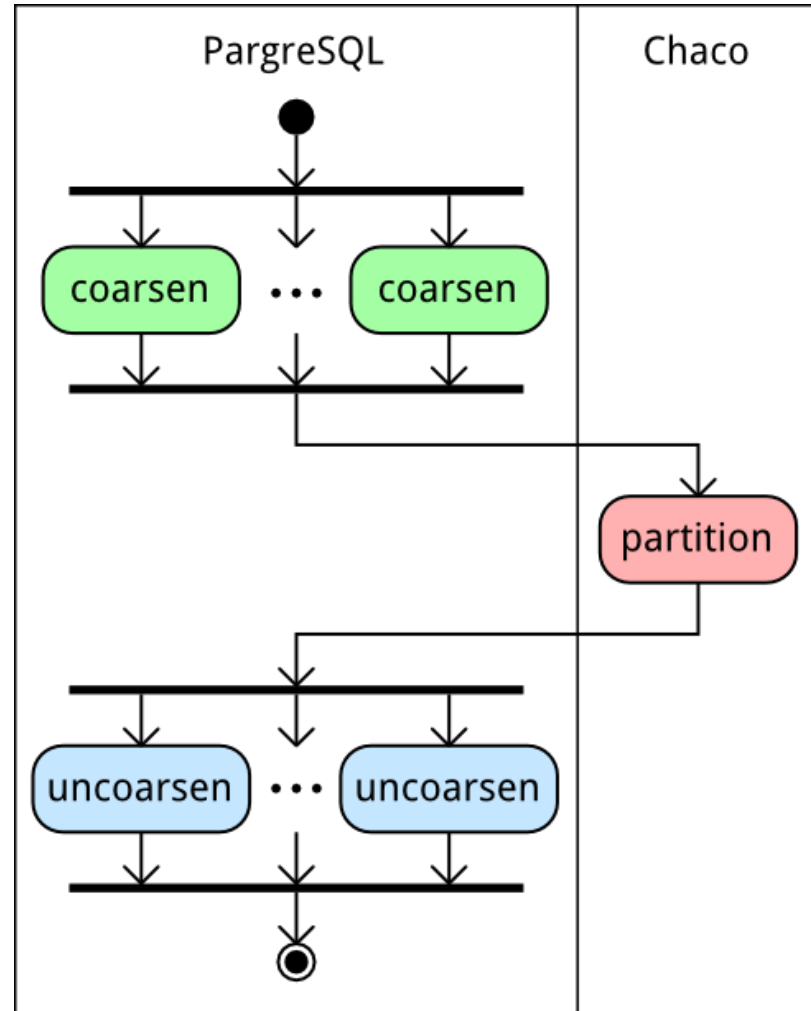
| Rank | Company | System   | Performance (tpmC) | DBMS   | OS                            |
|------|---------|--|--------------------|--|-------------------------------|
| 1    | Oracle  | SPARC SuperCluster with T3-4 Servers             | 30 249 688         | Oracle Database 11g R2 Enterprise Edition w/RAC w/Partitioning | Oracle Solaris 10 09/10       |
| 2    | IBM     | IBM Power 780 Server Model 9179-MHB              | 10 366 254         | IBM DB2 9.7  | AIX Version 6.1               |
| 3    | Oracle  | Sun SPARC Enterprise T5440 Server Cluster        | 7 646 486          | Oracle Database 11g Enterprise Edition w/RAC w/Partitioning    | Sun Solaris 10 10/09          |
|      | SUSU    | Tornado  | 2 202 531          | PargreSQL  | Linux CentOS 6.2              |
| 4    | HP      | HP Integrity rx5670 Cluster Itanium2/1.5 GHz-64p | 1 184 893          | Oracle Database 10g Enterprise Edition                         | Red Hat Enterprise Linux AS 3 |

# Эксперименты: разбиение графов

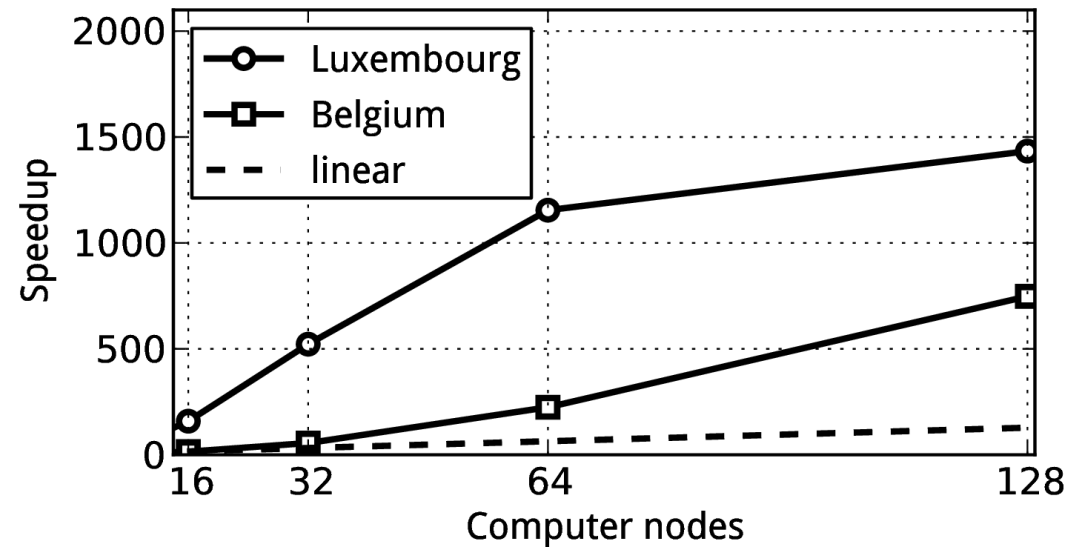
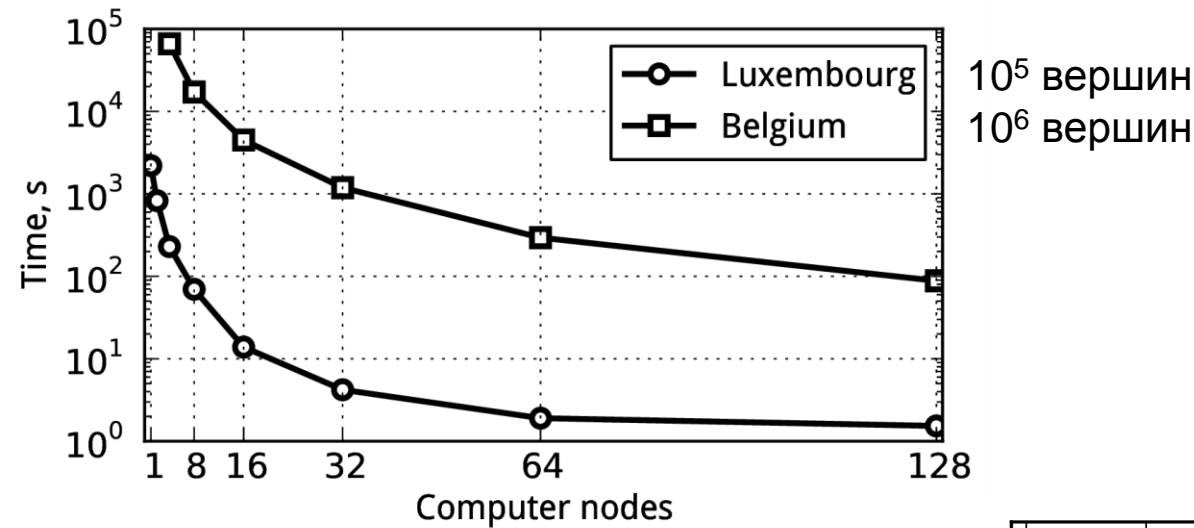


cut size  $\rightarrow$  min

○ partition size  $\approx$  ● partition size



# Эксперименты: разбиение графов





# Заключение

- ❑ Параллельная СУБД PargreSQL:  
внедрение фрагментного параллелизма  
в свободную СУБД PostgreSQL
  - ❑ Эксперименты:  
приемлемая эффективность PargreSQL  
на задачах обработки сверхбольших данных
  - ❑ Внедрение параллелизма в др. свободные СУБД  
(MySQL и проч.) – это возможно!
- 
- ❑ Спасибо за внимание
    - Вопросы?
    - Михаил Цымблер [zymbler@gmail.com](mailto:zymbler@gmail.com)
    - Константин Пан [kvapen@gmail.com](mailto:kvapen@gmail.com)