



Интеграция параллелизма в СУБД с открытым кодом

Л.Б. Соколинский

проректор по информатизации
ЮУрГУ, Челябинск

доктор физ.-мат. наук, профессор

sokolinsky@acm.org

Конференция «Большие Данные в национальной
экономике», 22.10.2013



PostgreSQL

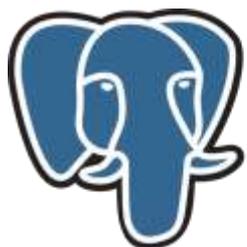
- Свободная объектно-реляционная СУБД с открытым исходным кодом для различных ОС
- Разрабатывается группой энтузиастов с 1995 г. (ответвление проекта POSTGRES М. Стоунбрейкера)
- Основные особенности
 - Поддержка SQL:2011 и ACID-транзакций
 - Хранимые процедуры SQL, pgSQL, Perl, Python, C и др.
 - Макс размеры: таблица – 32 Тб, поле – 1 Гб
 - Качественная документация исходного кода



PostgreSQL

- Надежная альтернатива коммерческим СУБД
- Примеры применения (<http://postgresql.org/about/users/>)
 - Корпорации
 - Apple
 - Sun
 - Cisco
 - Fujitsu
 - Red Hat
 - ...
 - Правительственные организации
 - U.S. State Department
 - United Nations Industrial Development Organization
 - ...

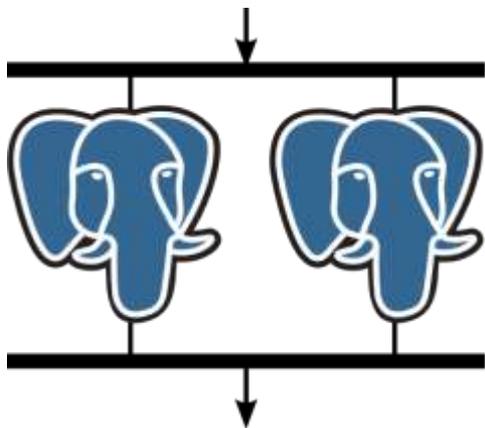
Проект PargreSQL



PostgreSQL

+

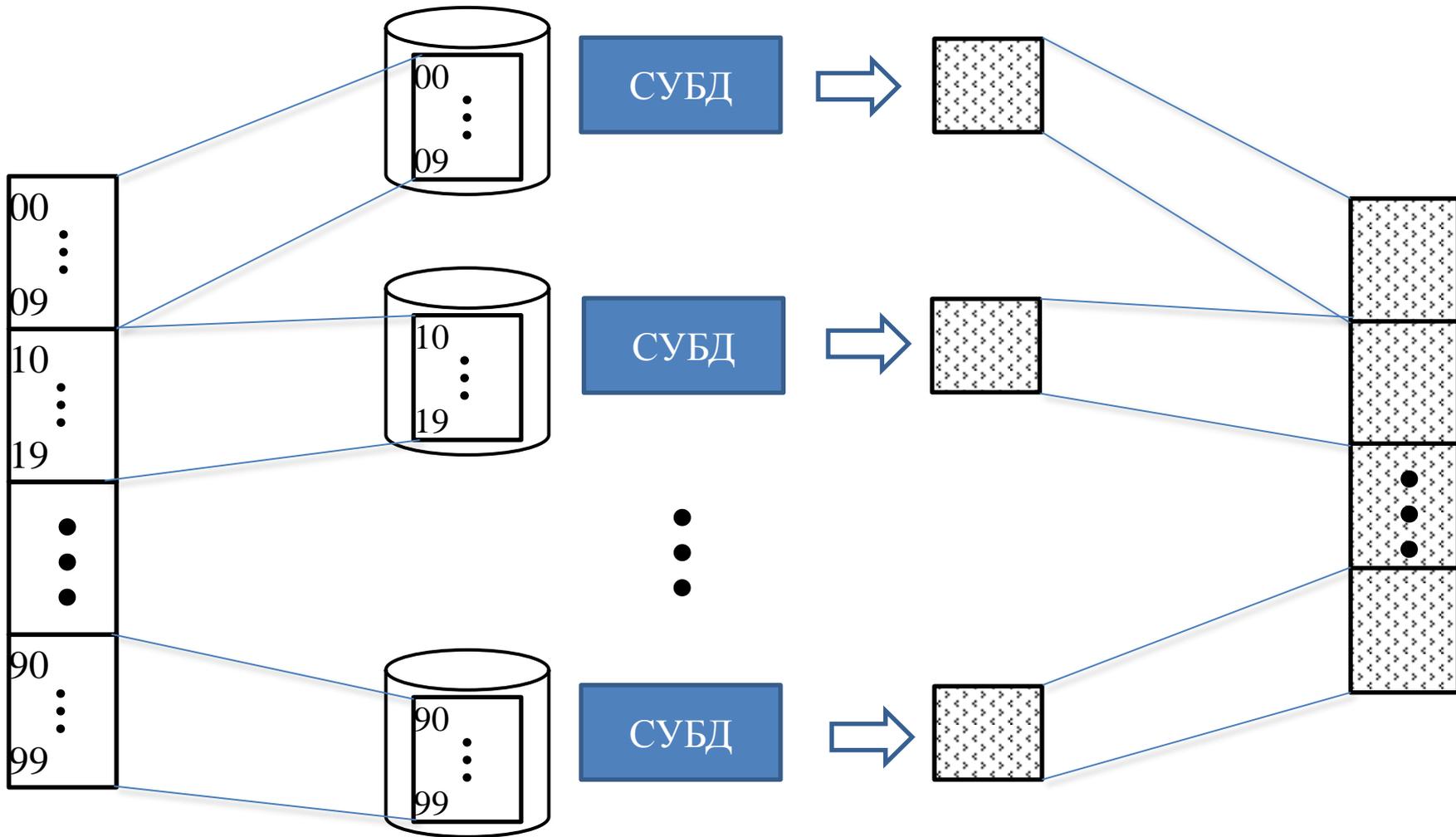
**фрагментный
параллелизм**



PargreSQL

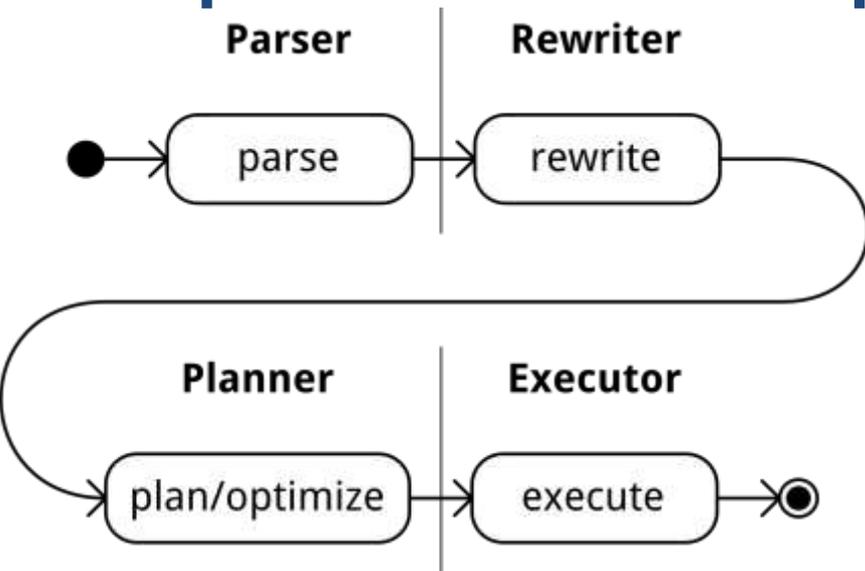


Фрагментный параллелизм

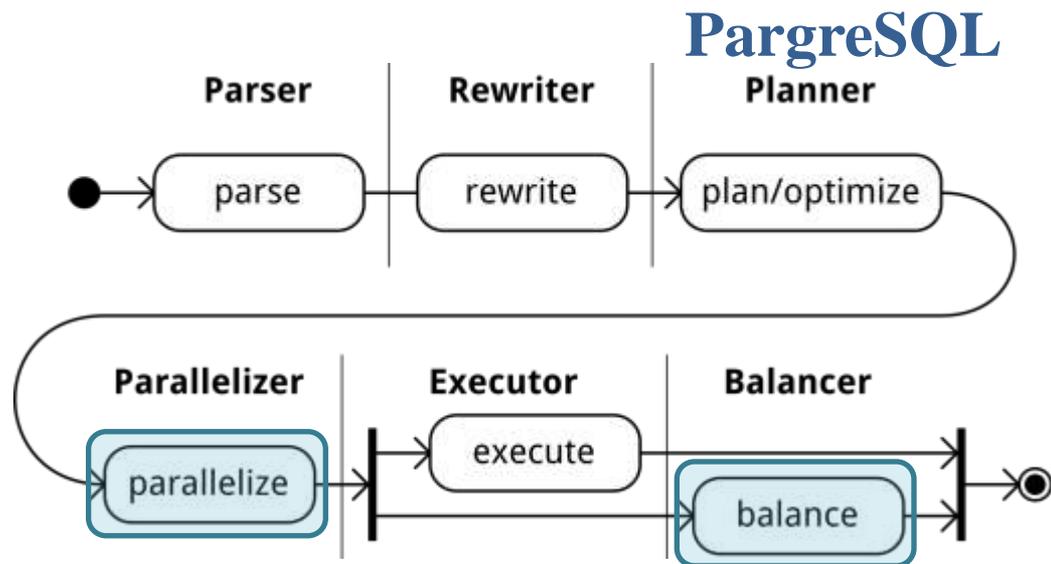




Обработка запроса



PostgreSQL



PargreSQL



Фрагментация таблиц

```
create table T (  
    ID integer primary key,  
    Column1 ...  
    Column2 ...)
```

```
with (fragattr = ID);
```

- Атрибут фрагментации: ID
- Функция фрагментации: $ID \bmod N$
- (N – количество узлов кластера)

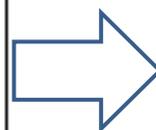
Совместимость Postgres→Pargres

par_Compat

```
#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)
```

```
// 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>

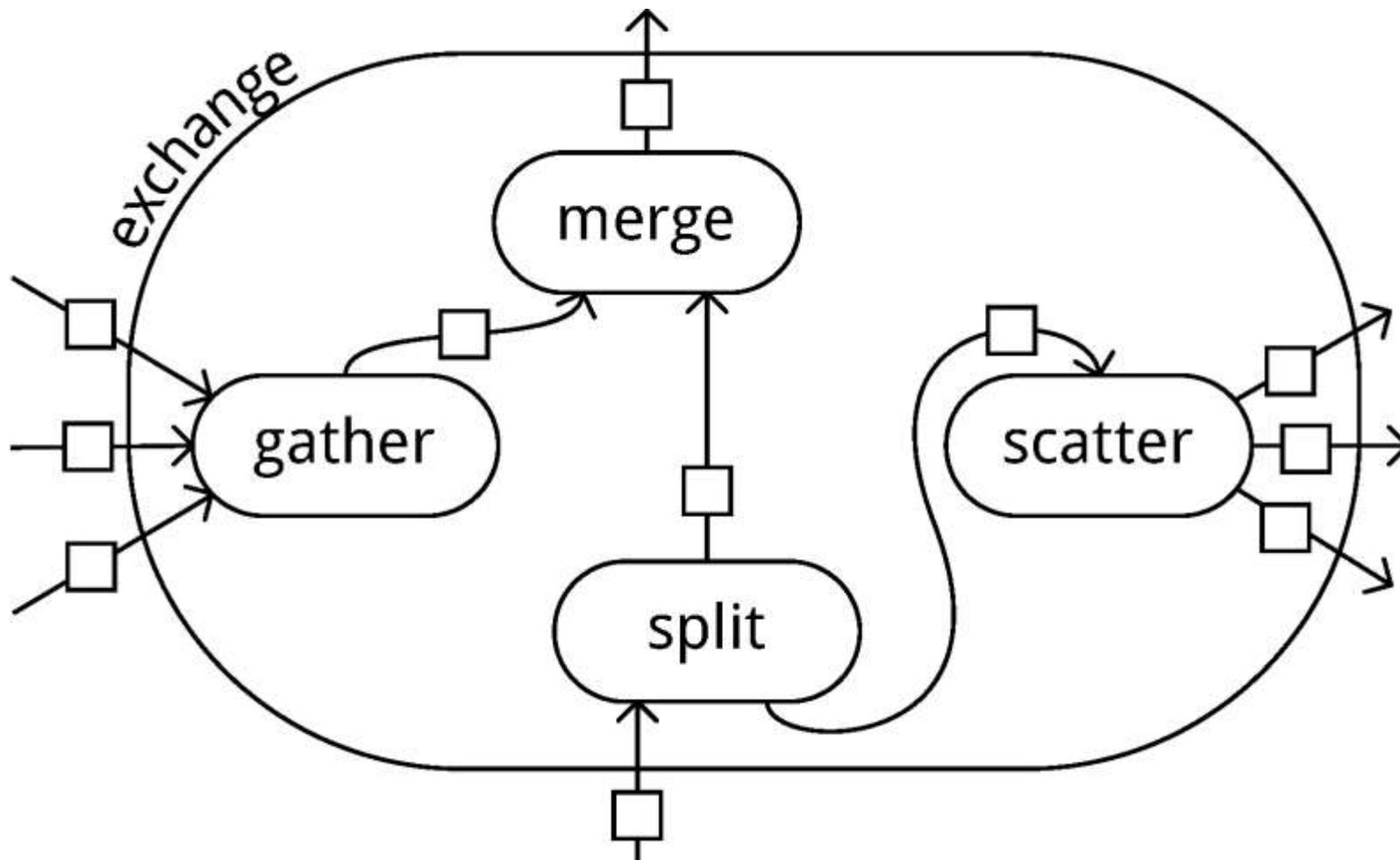
void main()
{
    PGconn c = PQconnectdb(...);
    PGresult r = PQexec(c, ...);
    ...
    PQfinish(c);
}
```

PostgreSQL приложение

PargresSQL приложение



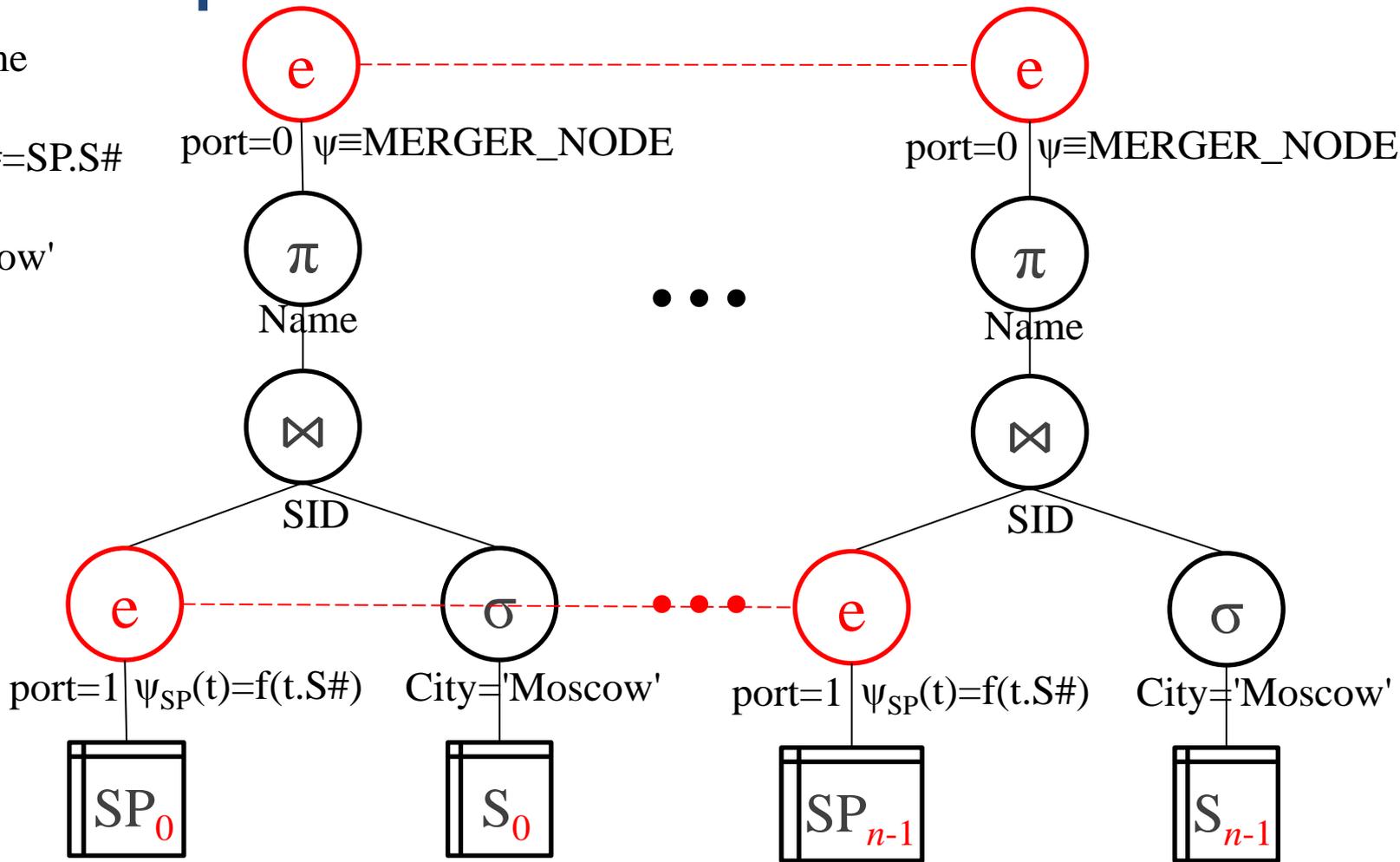
Оператор EXCHANGE





Оператор EXCHANGE

SELECT Name
FROM S, SP
WHERE S.S#=SP.S#
and
S.City='Moscow'



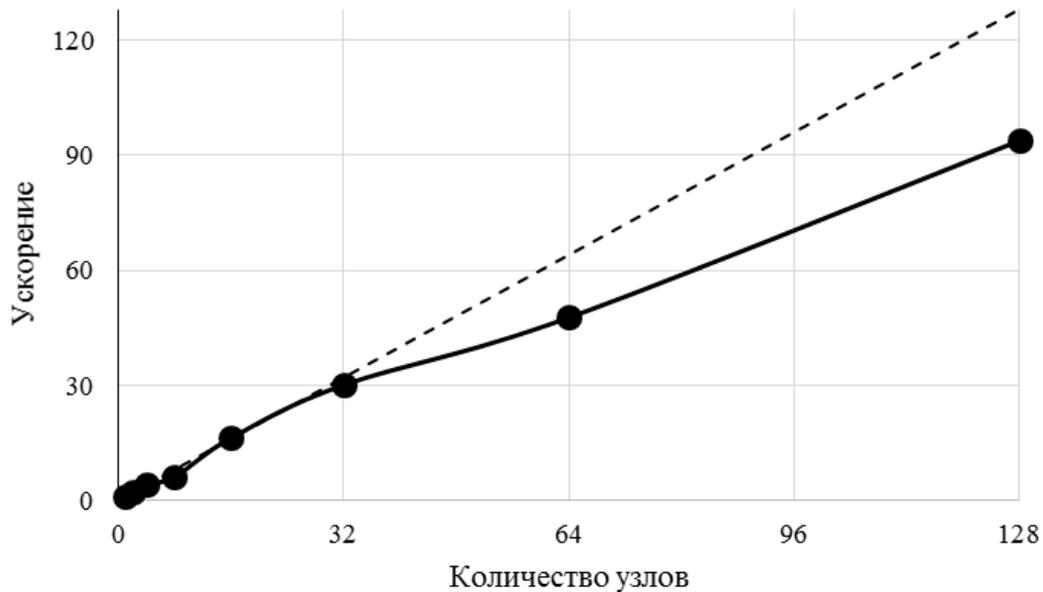
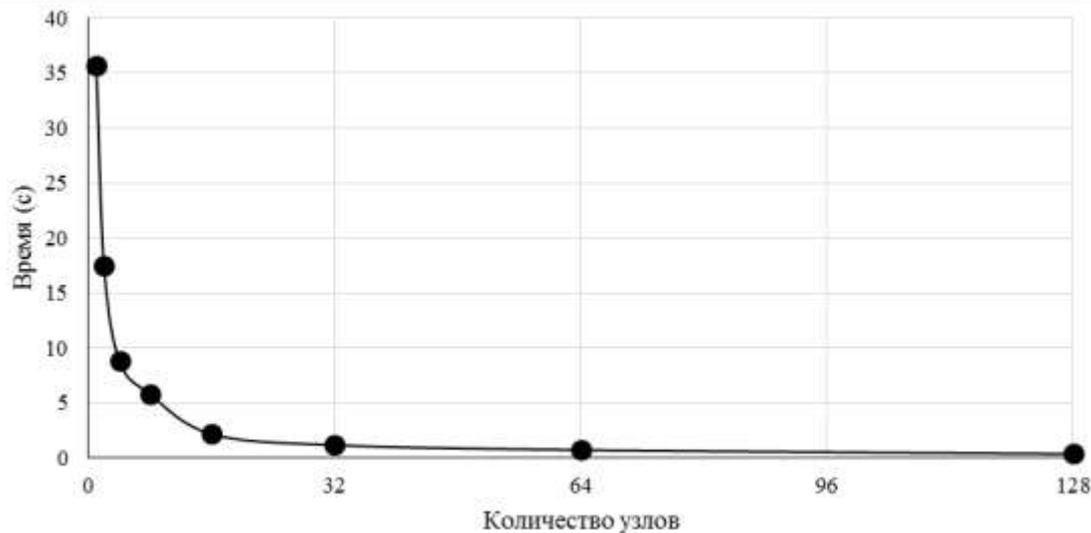


Эксперименты

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



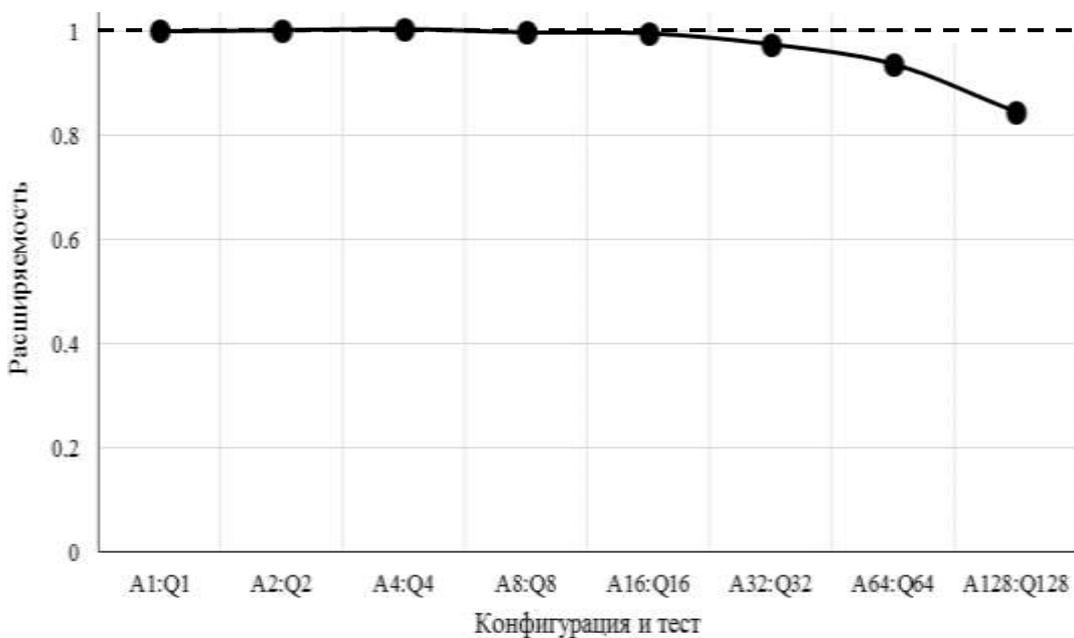
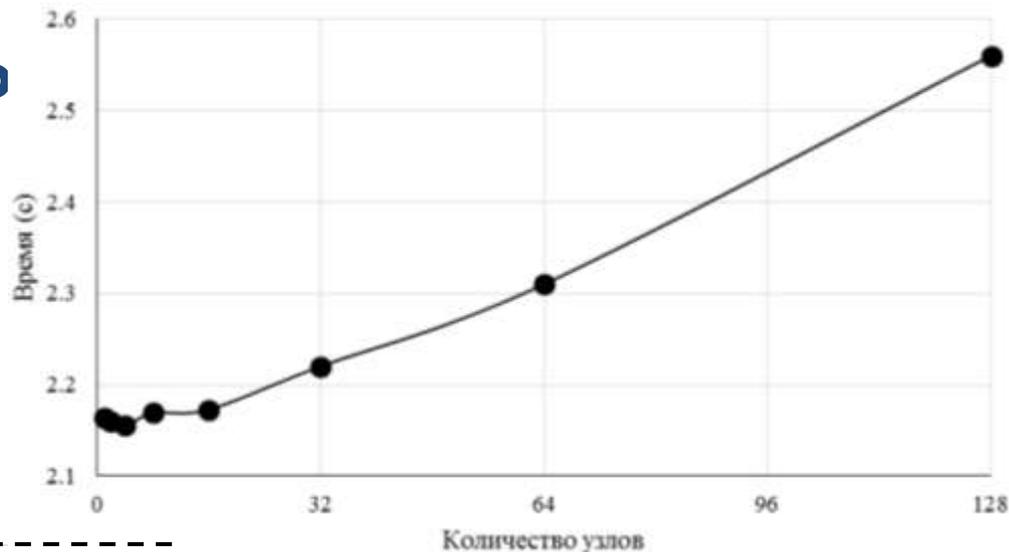
Ускорение



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



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



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?
 - Леонид Борисович Соколинский
sokolinsky@acm.org