

Конференция

«Системы хранения данных: максимум возможностей»  
(22 марта 2011 г., Челябинск)

# Организация систем хранения данных на базе вычислительных кластеров

М.Л. Цымблер, Л.Б. Соколинский

Южно-Уральский государственный университет (Челябинск)

# Содержание

- Вычислительные мощности Лаборатории суперкомпьютерного моделирования (ЛСМ)
  - СКИФ Урал
  - СКИФ Аврора
- Проекты ЛСМ, с хранением и обработкой данных сверхбольших объемов
  - PargreSQL
  - MedMining
  - Energo
  - VideoStorm

# Динамика развития суперкомпьютерных мощностей в ЮУрГУ

Вычислительный кластер  
**Infinity**  
Пиковая производительность  
**333 Gflops**  
**2004 г.**



**Intel Xeon64**  
**DP 3,2 ГГц**

Вычислительный кластер  
**СКИФ Урал**  
Пиковая производительность  
**16 Teraflops**  
**2008 г.**



**Intel Xeon E5472**  
**3 ГГц**

Суперкомпьютер  
**СКИФ-Аврора ЮУрГУ**  
Пиковая производительность  
**24 Teraflops**  
**2010 г.**



**Intel Xeon x5570**  
**2,93 ГГц**

Суперкомпьютер  
**СКИФ-Аврора ЮУрГУ**  
Пиковая производительность  
**117.6 Teraflops**  
**2010 г.**



**Intel Xeon x5680**  
**3,33 ГГц**

# Кластер СКИФ Урал

- Число вычислительных узлов/процессоров/ядер: **166/332/1328**
- Тип процессора: **Intel Xeon E5472 (4 ядра по 3.0 GHz)**
- Коммуникационные сети:
  - Системная: **InfiniBand (20 Gbit/s, макс. задержка 2  $\mu$ s)**
  - Управляющая: **Gigabit Ethernet**
  - Сервисная: **СКИФСervNet**
- Оперативная память: **1.33 TB**
- Дисковая память: **49.29 TB**
- Параллельная система хранения данных: **T-Platforms ReadyStorage – 20TB**



# Суперкомпьютер СКИФ Аврора

- **4** планируемое место в рейтинге СНГ TOP50
- **70** планируемое место в рейтинге TOP500
- **117,64** TFlops
- **8832** вычислительных ядер
- Коммуникационные сети:
  - Трехмерный топ: **60** Гбит/сек
  - InfiniBand QDR: **40** Гбит/сек
  - Сети мониторинга: **3** шт.
- Оперативная память: **8.8** Тбайт
- Дисковая память: **108** Тбайт



# Параллельная система хранения

- Panasas ActiveStorage 5100 20Тб
  - Производительное хранилище данных с архитектурой NAS, разработанное специально для кластеров
  - Пропускная способность более 60 Гбит/сек
  - Параллельный доступ к данным всех узлов кластера
  - Рост объема системы не усложняет управляемость

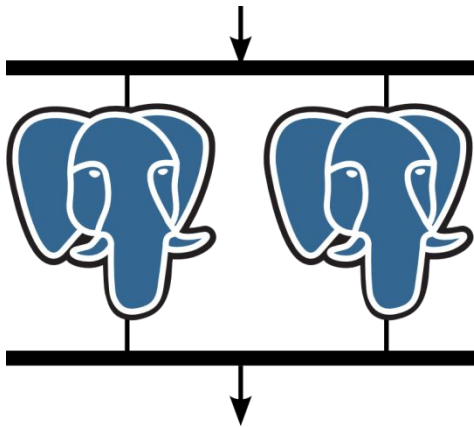


# Проект PargreSQL

- Внедрение параллелизма в свободную последовательную СУБД PostgreSQL



PostgreSQL

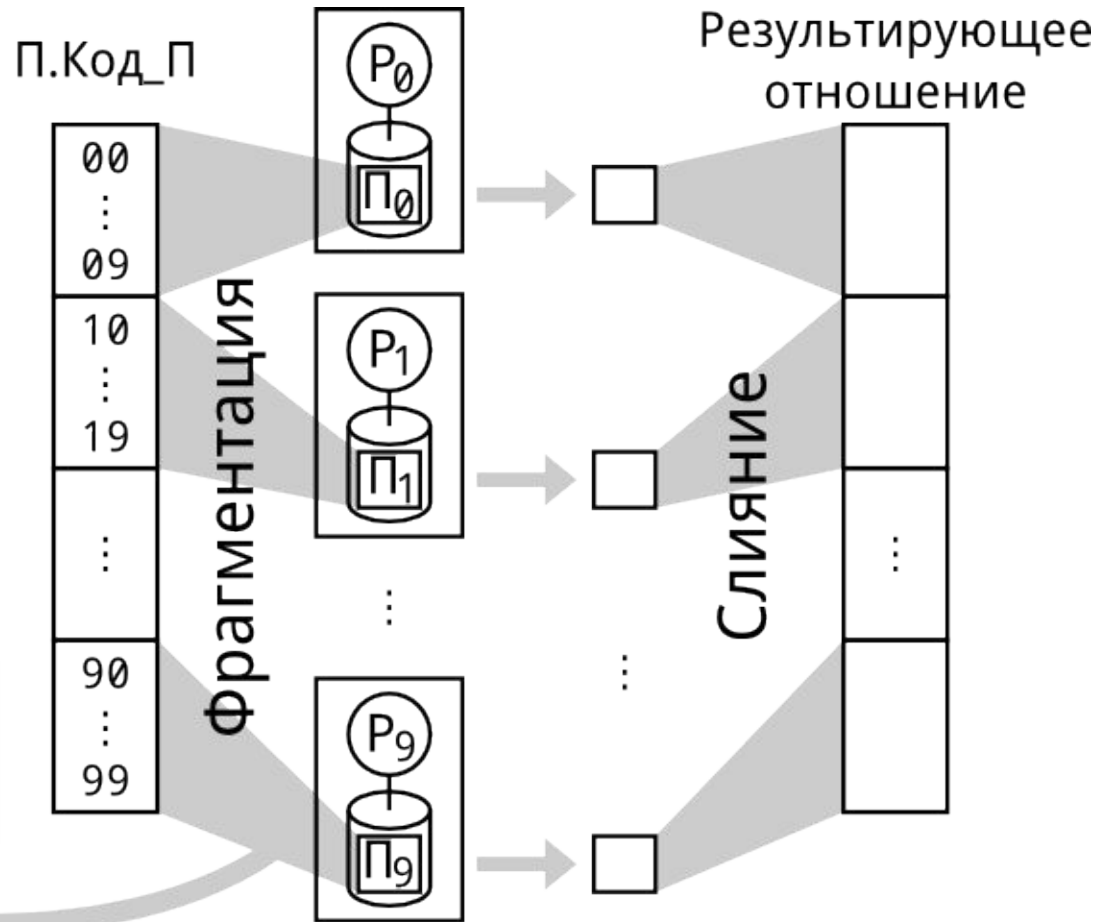


PargreSQL

# Фрагментный параллелизм

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

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

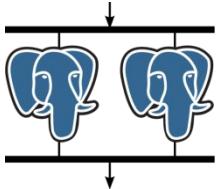
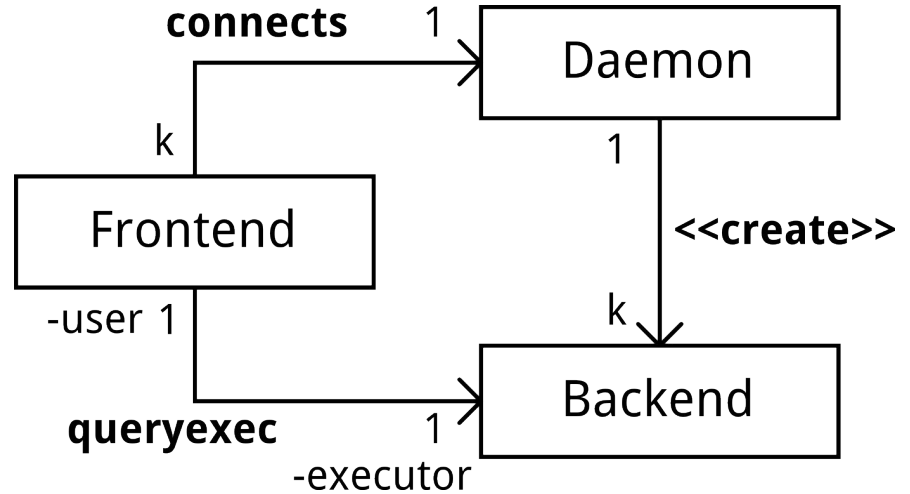




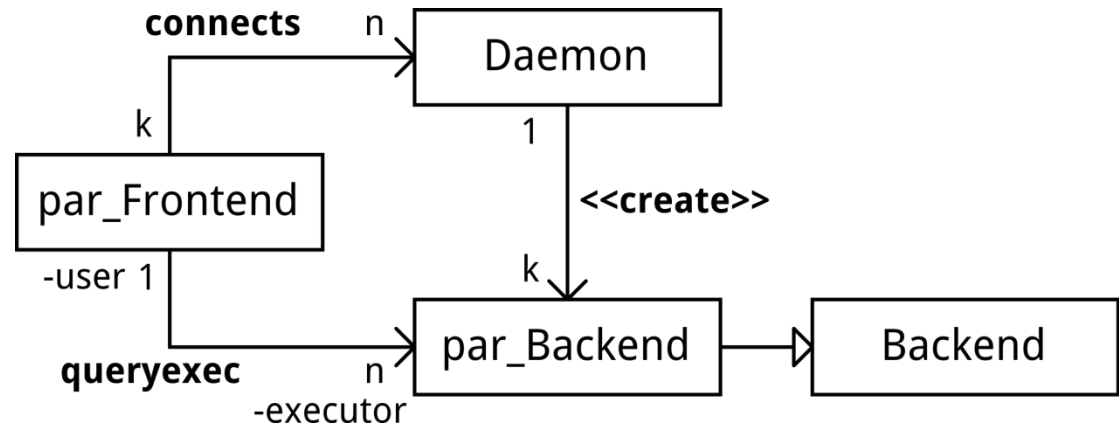
# PostgreSQL vs PargreSQL



PostgreSQL



PargreSQL



# Проект MedMining



- Параллельная система сбора, накопления и анализа данных по эффективному использованию ресурсов человека при предельных физических нагрузках
- Аппаратно-программный комплекс Oхусon Mobile
  - Более 40 различных показателей функционирования организма человека в режиме реального времени с дискретизацией до 1 сек.
  - Прирост базы данных исследований: 1 Гб в день

# Проект Energo



- Создание автоматизированного ситуационного центра энергоэффективности
  - Прирост базы данных показателей: 1 Гб в месяц
- Мониторинг и оптимизация потребления энергоресурсов объектов ЖКХ ЮУрГУ на основе оперативного и интеллектуального анализа данных об энергопотреблении

# Проект VideoStorm



- Разработка параллельной системы управления видеоархивом телеканала ЮУрГУ-TV
- Хранение сырого, монтированного и сжатого видео
- Прирост объема видеоархива: 30 Тб в год
- Эффективное хранение и обработка видео
  - сжатие видео
  - автоматическое разбиение на сцены
  - **высокая ГОТОВНОСТЬ ДАННЫХ**



# Спасибо за внимание

- Вопросы?

- Михаил Леонидович Цымблер

- доцент кафедры системного программирования ЮУрГУ, к.ф.-м.н.
- [mzym@susu.ru](mailto:mzym@susu.ru)