

Международная научная конференция
"Суперкомпьютерные системы и их применение"

26-28 октября 2004 г., Минск (Республика Беларусь)

**Прототипирование параллельной СУБД
как основа учебного курса
по параллельным системам баз данных***

М.Л. Цымблер, Л.Б. Соколинский, А.В. Лепихов
Южно-Уральский государственный университет

* Работа выполнена при финансовой поддержке РФФИ, проект 03-07-90031

Мотивация и цели

■ *Мотивы*

- Актуальность подготовки в вузе специалистов в области параллельных систем баз данных.
- Фактическое отсутствие российских и зарубежных образовательных Интернет-ресурсов по параллельным системам баз данных.

■ *Цель*

- Разработка электронного учебного курса по параллельным системам баз данных.

Электронный учебный курс

Курс "Параллельные системы баз данных" - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://pdbs.susu.ru/> Переход Links


Параллельные системы баз данных

Л.Б. Соколинский, М.Л. Цымблер

Электронный учебный курс
(учебное пособие)

Челябинск: ЮУрГУ, 2004

Sponsored by intel



[Аннотация](#)

[Программа курса](#)

[Тексты лекций](#)

[Слайды презентации](#)

[Задания к лабораторным работам](#)

[Внутренние спецификации](#)

[Описание комплексных тестов](#)

[Автономные тесты](#)

[Поддерживающая библиотека](#)

[Текст головного модуля](#)

[Файлы данных для комплексного тестирования](#)

[Исходные тексты поддерживающей библиотеки \(ограниченный доступ\)](#)

Аннотация

Содержание курса

Курс посвящен одному из самых сложных и интересных разделов системного программирования - параллельным системам баз данных. Параллельные системы баз данных призваны решать задачи, связанные с обработкой и поддержанием сверхбольших баз данных. Сверхбольшие базы данных характерны для таких приложений, как электронная коммерция, электронные библиотеки, геоинформационные системы, мультимедийные архивы, научные базы данных. Курс ориентирован на студентов старших курсов и аспирантов, специализирующихся по направлению "Информационные технологии".

Internet

Аудитория, уровень подготовки

■ *Аудитория*

- Студенты старших курсов и аспиранты – будущие специалисты в области системного программирования.

■ *Уровень подготовки*

- Знание материала университетского курса по системам баз данных.
- Навыки программирования на языке С.

Практикум – основа курса

- *Задача* – разработка *прототипа* параллельной СУБД, выполняющего
 - компиляцию запроса к модельной базе данных
 - формирование параллельного плана запроса
 - исполнение запроса.
- *Методическое обеспечение*
 - Библиотека прототипа
 - Справочник по функциям
 - Набор контрольных тестов

Библиотека прототипа

- *Библиотека прототипа* – объектный код функций прототипа, предварительно разработанного авторами.
- Прототип, разработанный авторами
 - Операционная система: MS Windows
 - Среда разработки: MS Visual C++
 - Обмен сообщениями: библиотека MPICH
 - Объем исходного текста: ≈ 7000 строк

Справочник по функциям

```
In file parallel.h:
extern int pl\_ParallelizeQuery
(treenode\_t* root)

Распараллелить запрос
```

Documentation

Данная функция выполняет формирование параллельного плана исполнения запрос атрибута операции EXCHANGE "номер порта" используются целые положительные создани параллельного агента.

Parameters:

root - указатель на корень дерева последовательного плана исполнения запроса

Returns:

0 в случае успеха или отрицательный код ошибки

See Also:

[treenode_t](#) [AgentManager](#) [am](#) [MakeAgent](#) [ErrorCodes](#)

[Alphabetic index](#)

Справочник по функциям прототипа параллельной СУБД

```
In file compiler.h:
typedef struct \_treenode treenode\_t

Узел дерева запроса
```

Documentation

Узел дерева запроса имеет следующие атрибуты:

N п/п	Атрибут	Семантика
1	operation_t operation	вид операции
2	void * cargo	пользовательское поле (сюда помещается дескриптор операции)
3	int fragmentatr	атрибут фрагментации поддерева, определяемого данным узлом
4	fragment_fn_t fragmentfn	функция фрагментации поддерева, определяемого данным узлом
5	int ary	арность (количество атрибутов) результирующего кортежа
6	execute_fn_t executefn	функция реализации операции в поддерева, определяемым данным узлом
7	reset_fn_t resetfn	функция реализации старта операции в поддерева, определяемым данным узлом
8	tuple_t result	указатель на кортеж-результат операции в поддерева, определяемым данным узлом
9	int isBOF	"начало отношения" (устанавливается в Истину перед тем, как скобочный шаблон обработал первый кортеж отношения)
10	int isBOF	"конец отношения" (устанавливается в Истину после того, как скобочный шаблон обработал последний кортеж отношения)
11	struct _treenode_t * left	указатель на левого сына
12	struct _treenode_t * right	указатель на правого сына.

See Also:

[operation_t](#) [opJOIN_t](#) [opRESTRICT_t](#) [opSCAN_t](#) [opMERGE_t](#) [opGATHER_t](#) [opSPLIT_t](#) [opSCATTER_t](#) [cp](#) [CompileQuery](#) [fragment_fn_t](#) [execute_fn_t](#)

[Alphabetic index](#)

Справочник по функциям прототипа параллельной СУБД

замечания и предложения направлять по адресу pladm@yandex.ru

Контрольный тест

- *Контрольный тест* – совокупность отношений модельной базы данных, запроса к ней на языке запросов прототипа, и результата выполнения запроса.

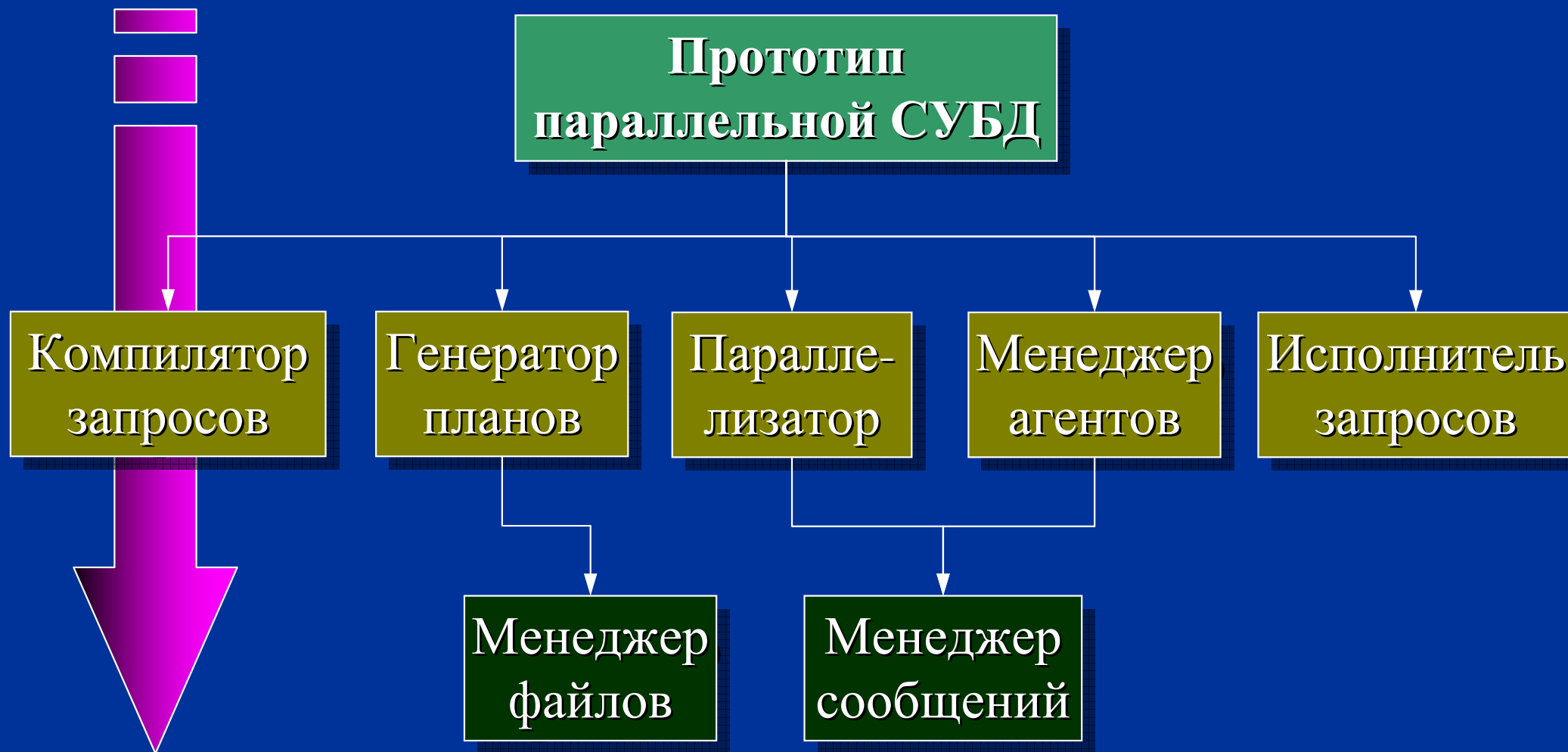
R0	A0	A1	A2	A3
F0	0	0	43	67
	1	0	21	22
	...			
F1	5	1	43	61
	6	1	6	23
	...			
F2	11	2	8	
	12	2	3	
	...			



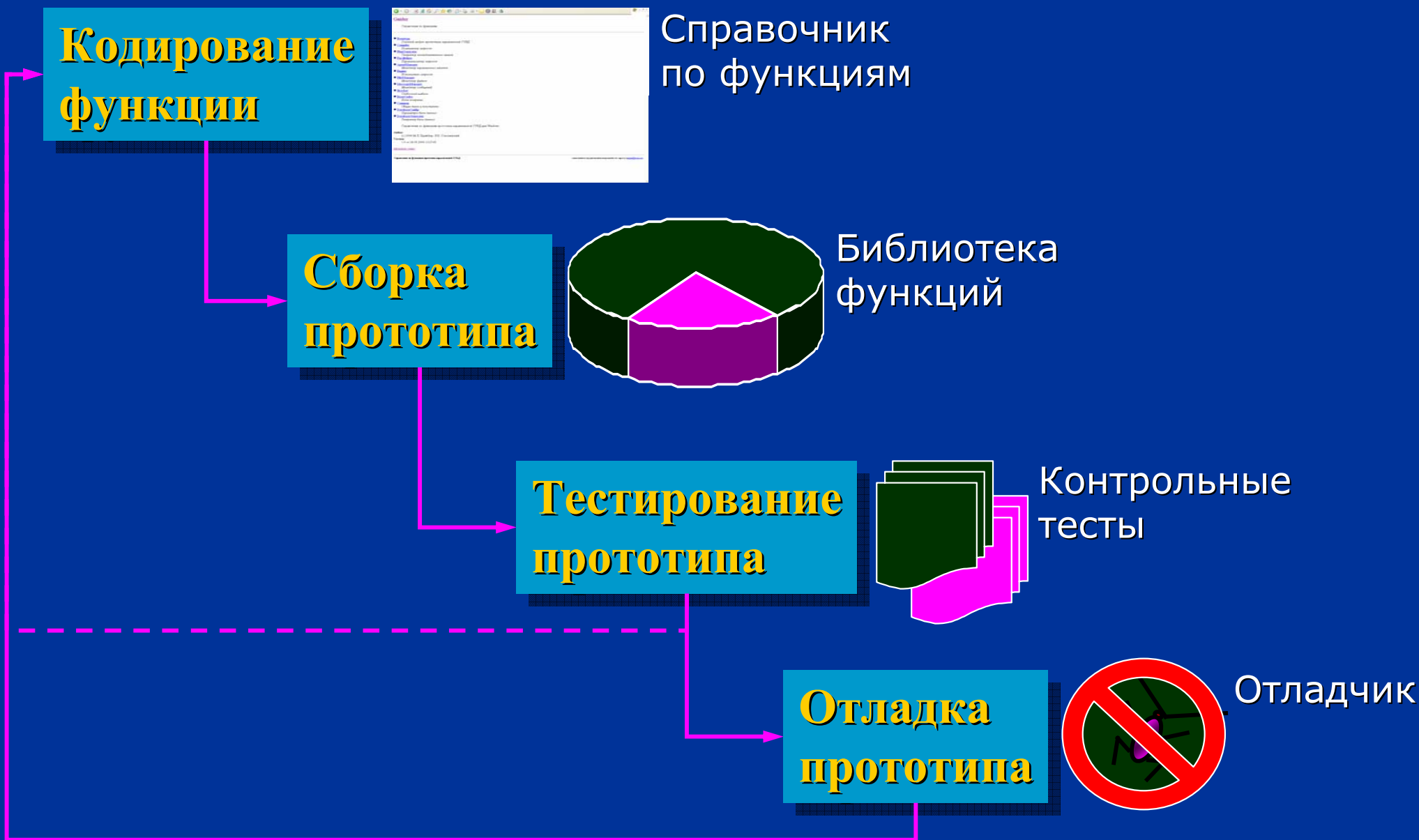
A0	A1	A2	A3	A4	A5	A6
2	0	80	71	100	21	22
4	0	80	58	100	21	22
8	1	80	24	107	21	28
11	2	80	67	113	21	3
11	2	80	67	116	21	23

(R0 where A2=80)
join
(R1 where A2=21)

Организация разработки



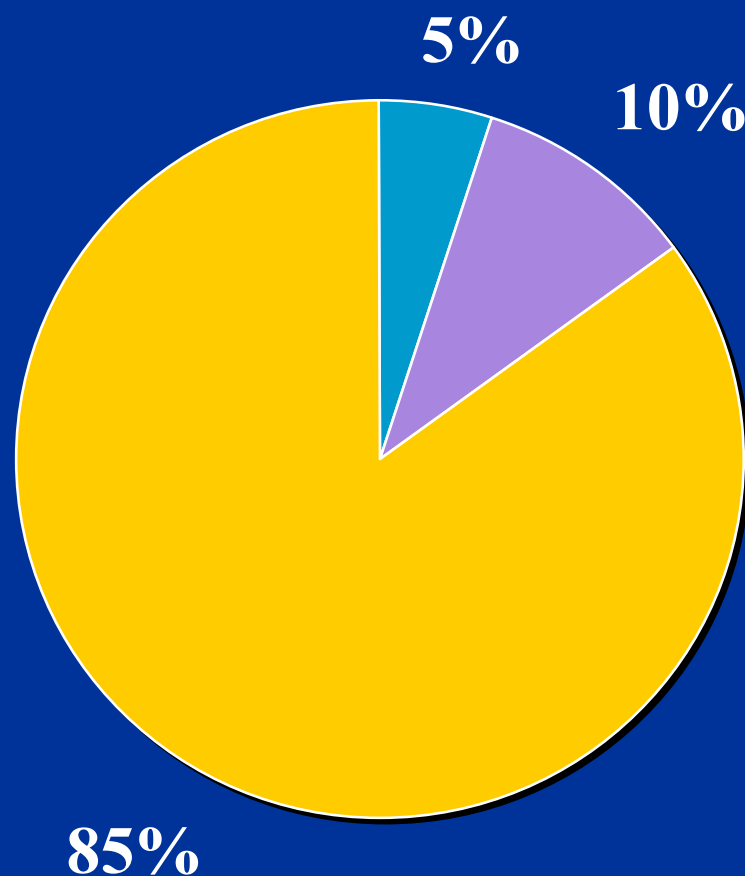
Цикл разработки



Апробация курса

- Зимняя школа-практикум молодых ученых и специалистов – 2004 (Нижегородский государственный университет).
- *Слушатели*
40 человек – студенты старших курсов, аспиранты и магистранты из России (МГУ, МФТИ, СПбГТУ, ННГУ, НГУ и др.), Беларуси (БГУ) и Украины (КНУ).
- *Распределение часов*
лекции – 6 часов, практикум – 12 часов.

Апробация: подходы к разработке

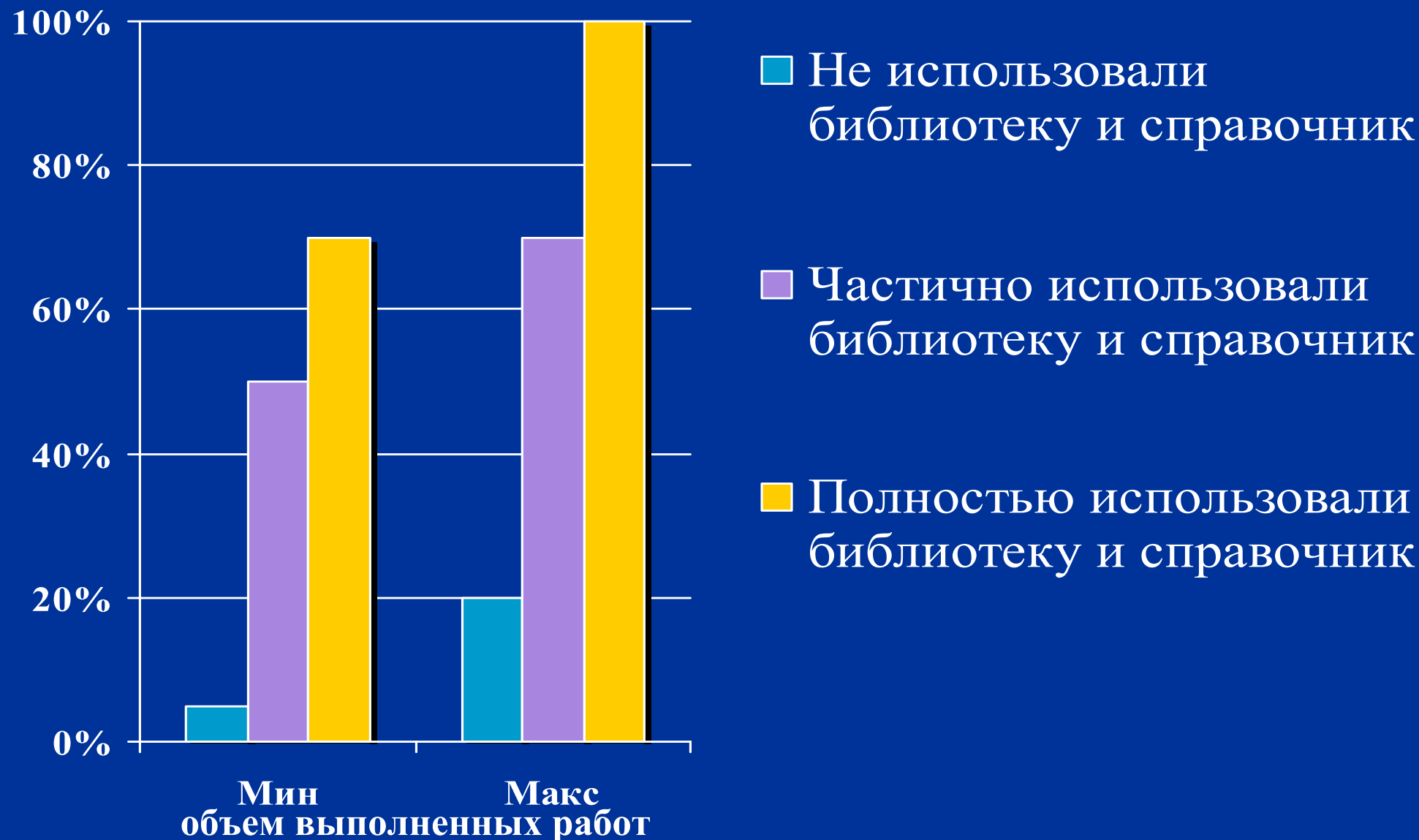


■ Не использовали библиотеку и справочник

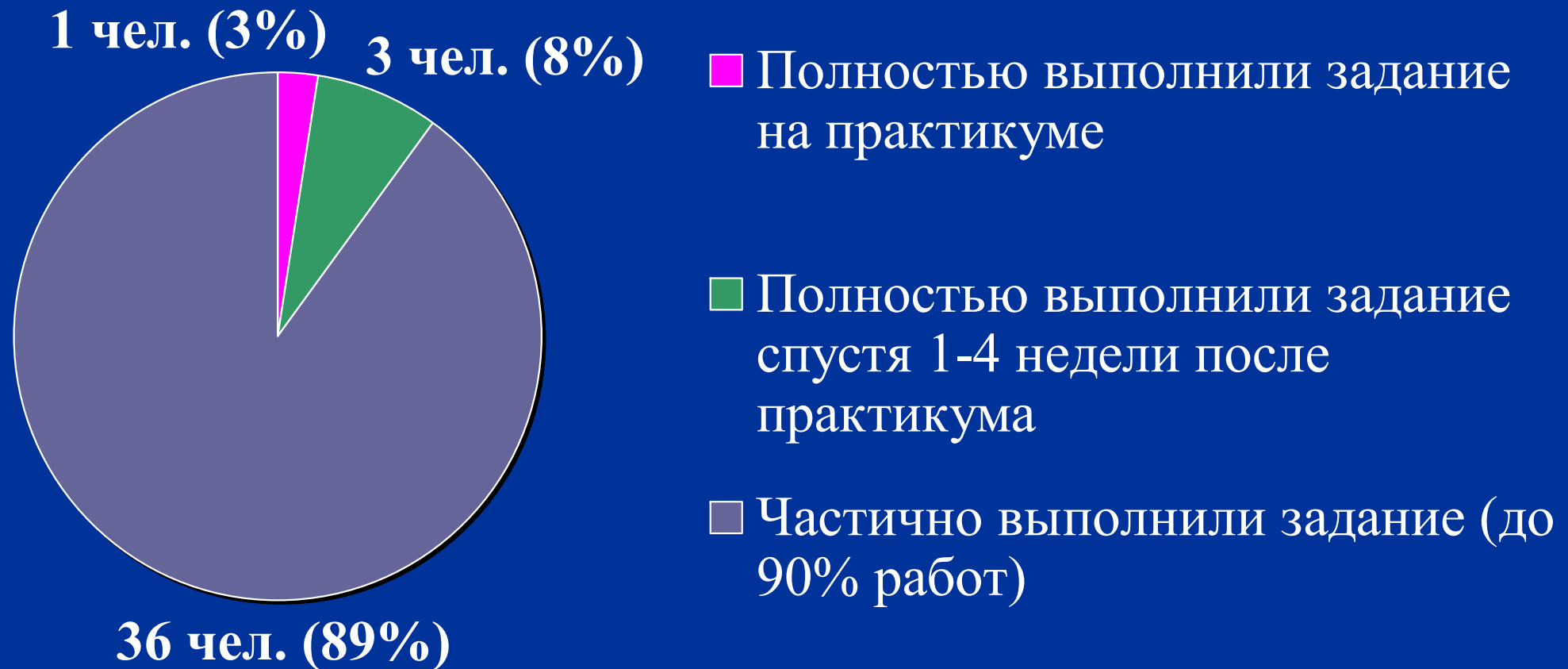
■ Частично использовали библиотеку и справочник

■ Полностью использовали библиотеку и справочник

Апробация: результаты разработки



Апробация: сложность разработки



Апробация: необходимость отладчика

- Для успешного выполнения заданий практикума необходим интерактивный отладчик
- Разработан эмулятор обменов сообщениями по стандарту MPI
 - Эмулятор – динамическая библиотека, в интерфейсе которой необходимые для разработки функции MPI.
 - Параллельный процесс – набор потоков Windows.
 - Прототип запускается непосредственно из среды MS Visual C++ (без загрузчика MPI-программ).
 - Переход от отладки к тестированию – смена параметров компоновки.

Заключение

- Рассмотрен подход к организации учебного курса по параллельным системам баз данных, практикум которого предполагает разработку прототипа параллельной СУБД.
- Данный подход успешно апробирован на Зимней школе-практикуме молодых ученых "Технологии параллельного программирования – 2004" (г. Нижний Новгород).
- Материалы курса размещены в сети Интернет:
<http://pdbs.susu.ru>