

Михаил Леонидович Цымблер

**Методы построения программного комплекса
для управления данными в вычислительных
системах с массовым параллелизмом**

**Диссертация
на соискание ученой степени кандидата физико-математических наук**

специальность 05.13.18

теоретические основы численного моделирования, численные методы и комплексы программ

Работа выполнена

на кафедре математического обеспечения ЭВМ
Челябинского государственного университета

Научный руководитель

зав. кафедрой математического обеспечения ЭВМ ЧелГУ
кандидат физ.-мат. наук, доцент Л.Б. Соколинский

Официальные оппоненты

зам. директора НИВЦ МГУ,
доктор физ.-мат. наук Вл.В. Воеводин

зав. отделом системного обеспечения ИММ УрО РАН
кандидат техн. наук, доцент В.В. Самофалов

Ведущая организация

Уральский государственный университет

Работа выполнена при финансовой поддержке РФФИ (проекты 97-07-90148, 00-07-90077)

Цель и задачи исследования

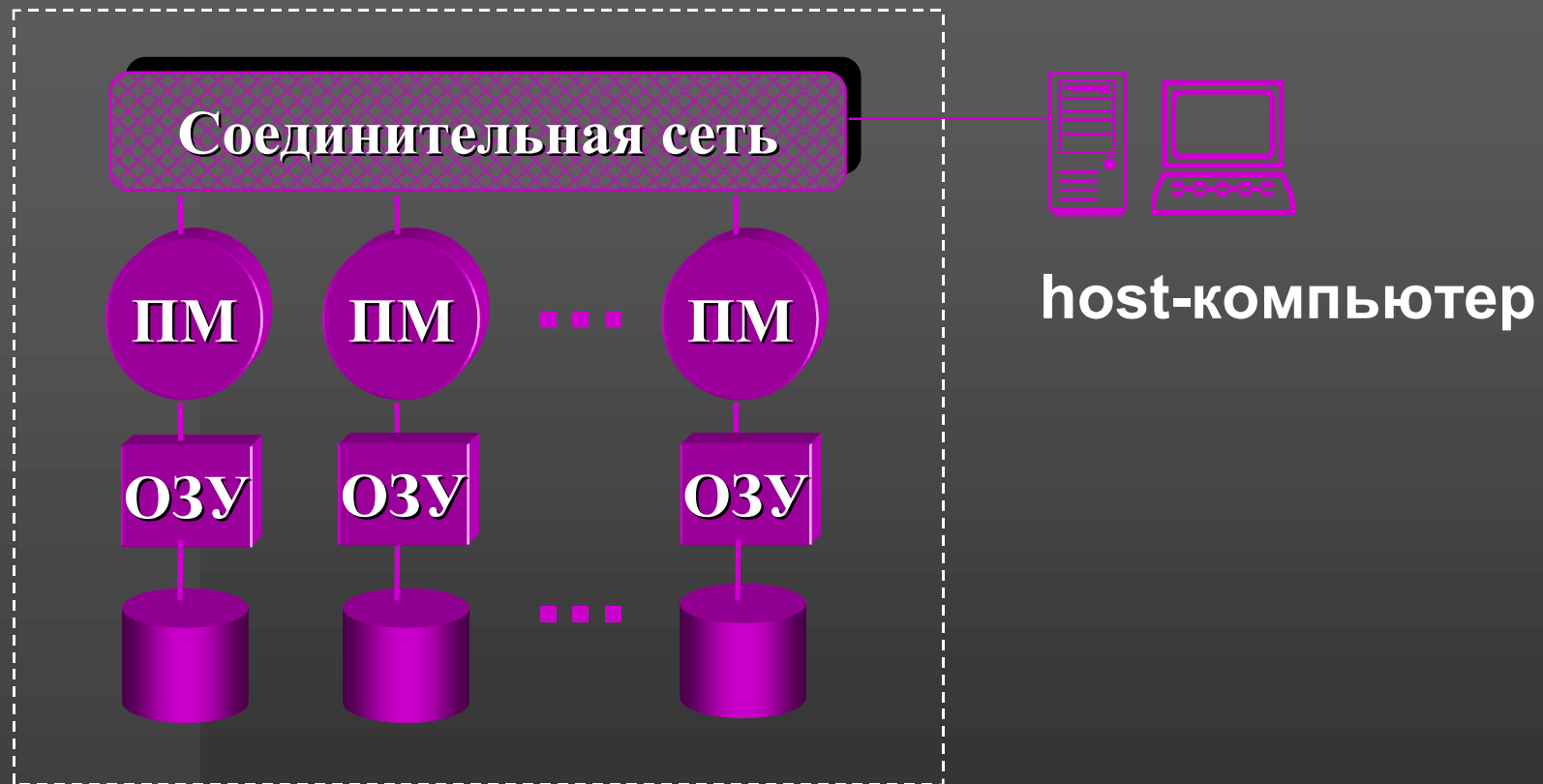
Цель

Исследование и разработка методов построения комплекса системных программ для управления данными в вычислительных системах с массовым параллелизмом

Задачи

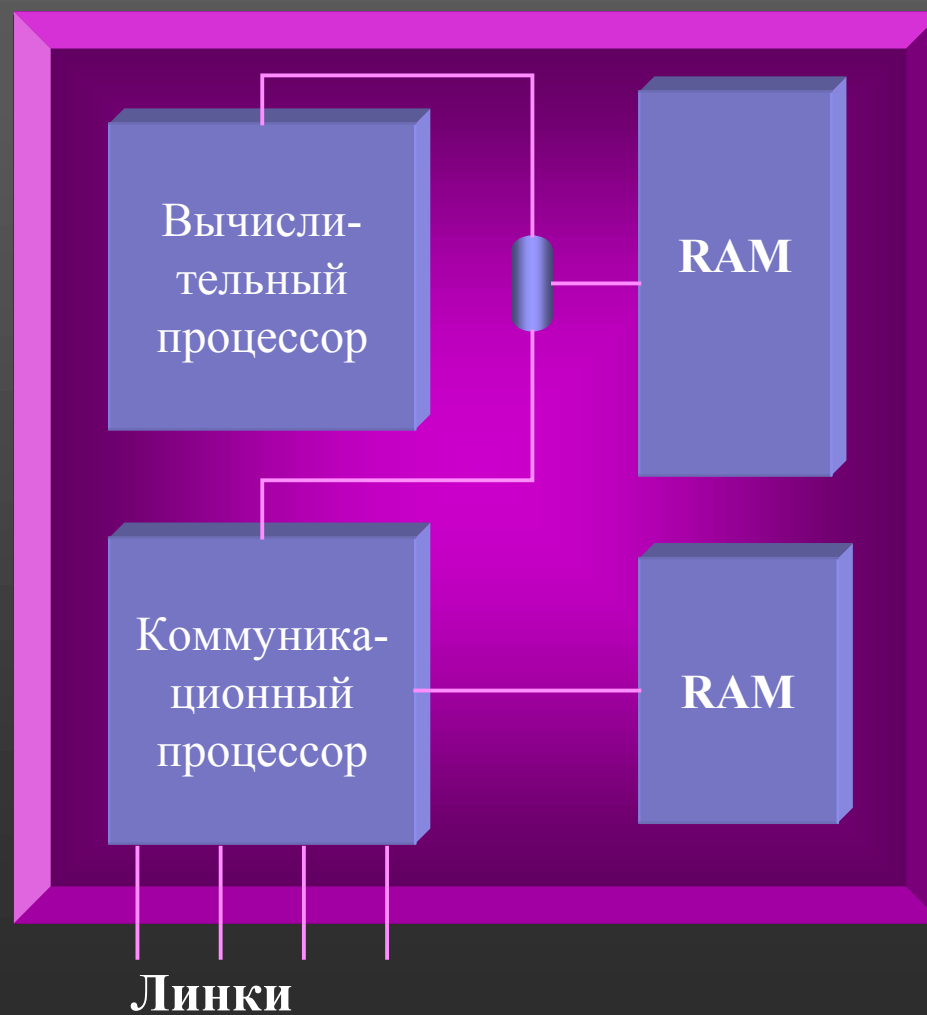
- исследование и анализ существующих методов организации хранения и передачи данных
- разработка новых методов управления данными, учитывающих особенности архитектуры современных многопроцессорных систем типа МВС-100/1000
- разработка программного комплекса для организации хранения и передачи данных в системе МВС-100/1000

Системы с массовым параллелизмом



ПМ – процессорный модуль

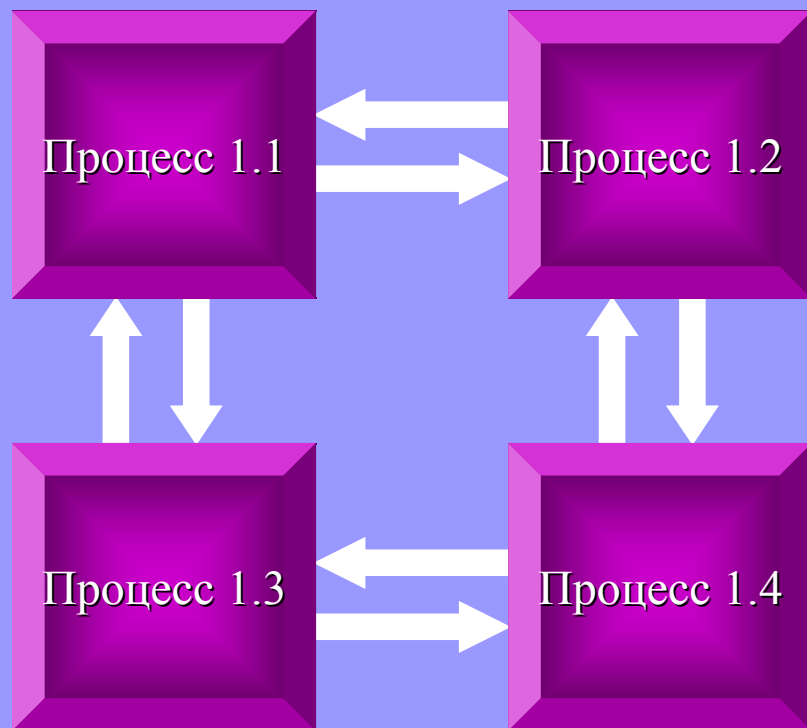
Типовой процессорный модуль МВС-100/1000



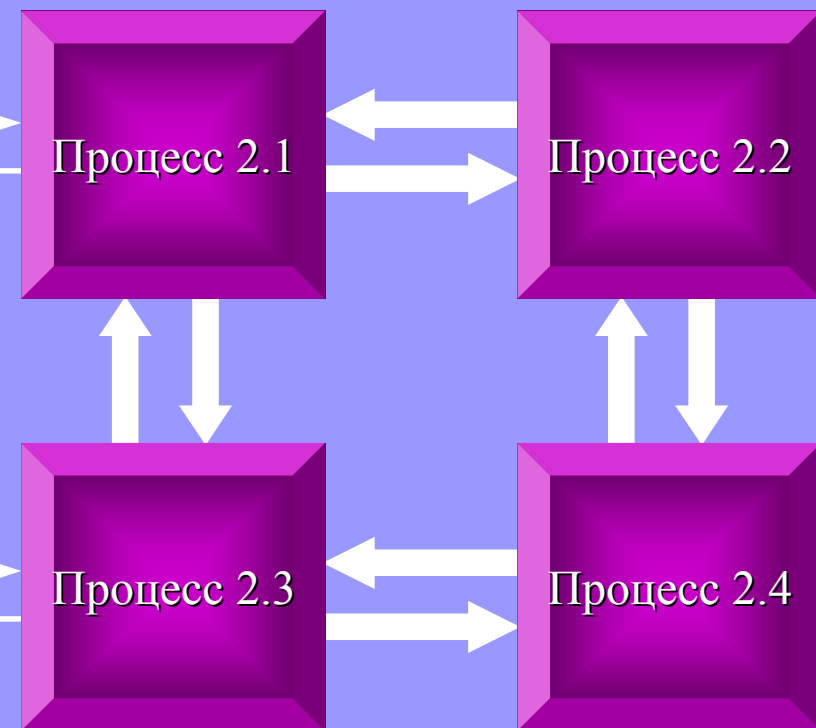
Анализ задач, решаемых на МВС

Задача

Подзадача 1

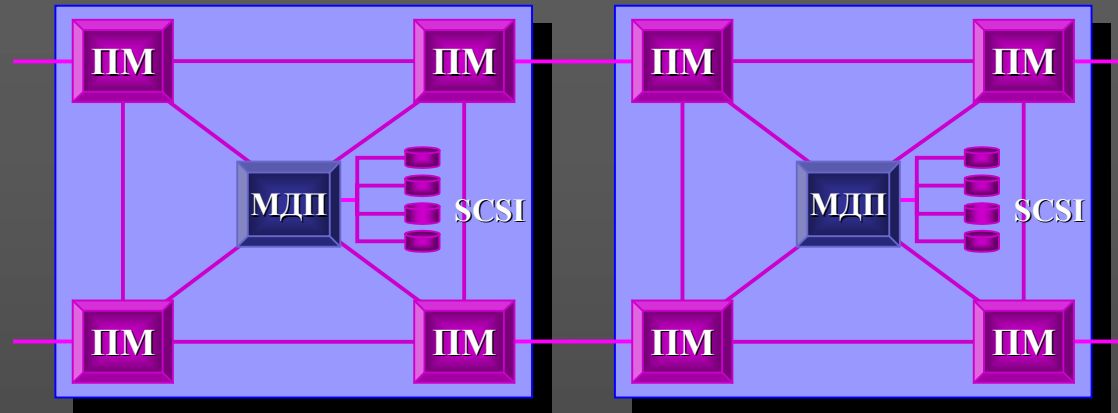


Подзадача 2

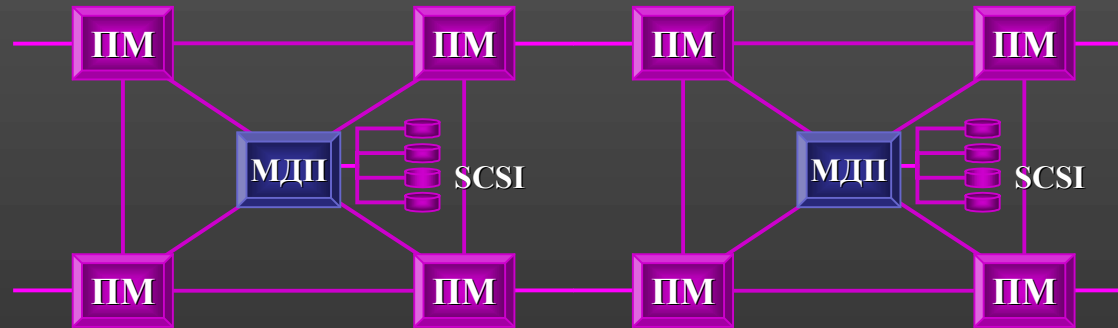


Иерархическая организация систем управления данными

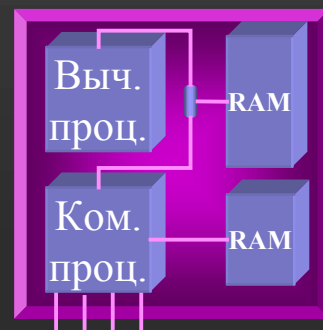
Логический уровень



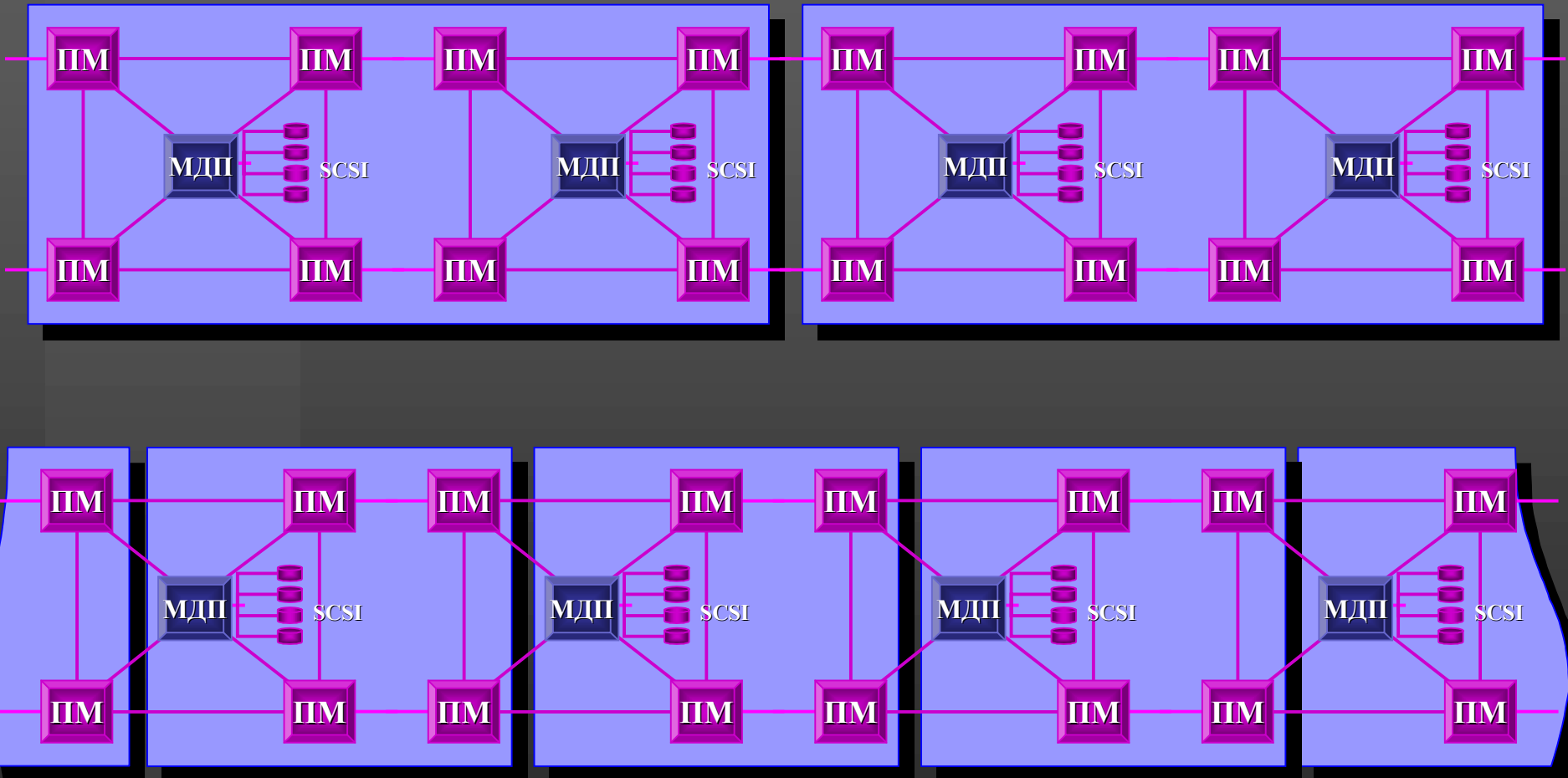
Физический уровень



Аппаратный уровень



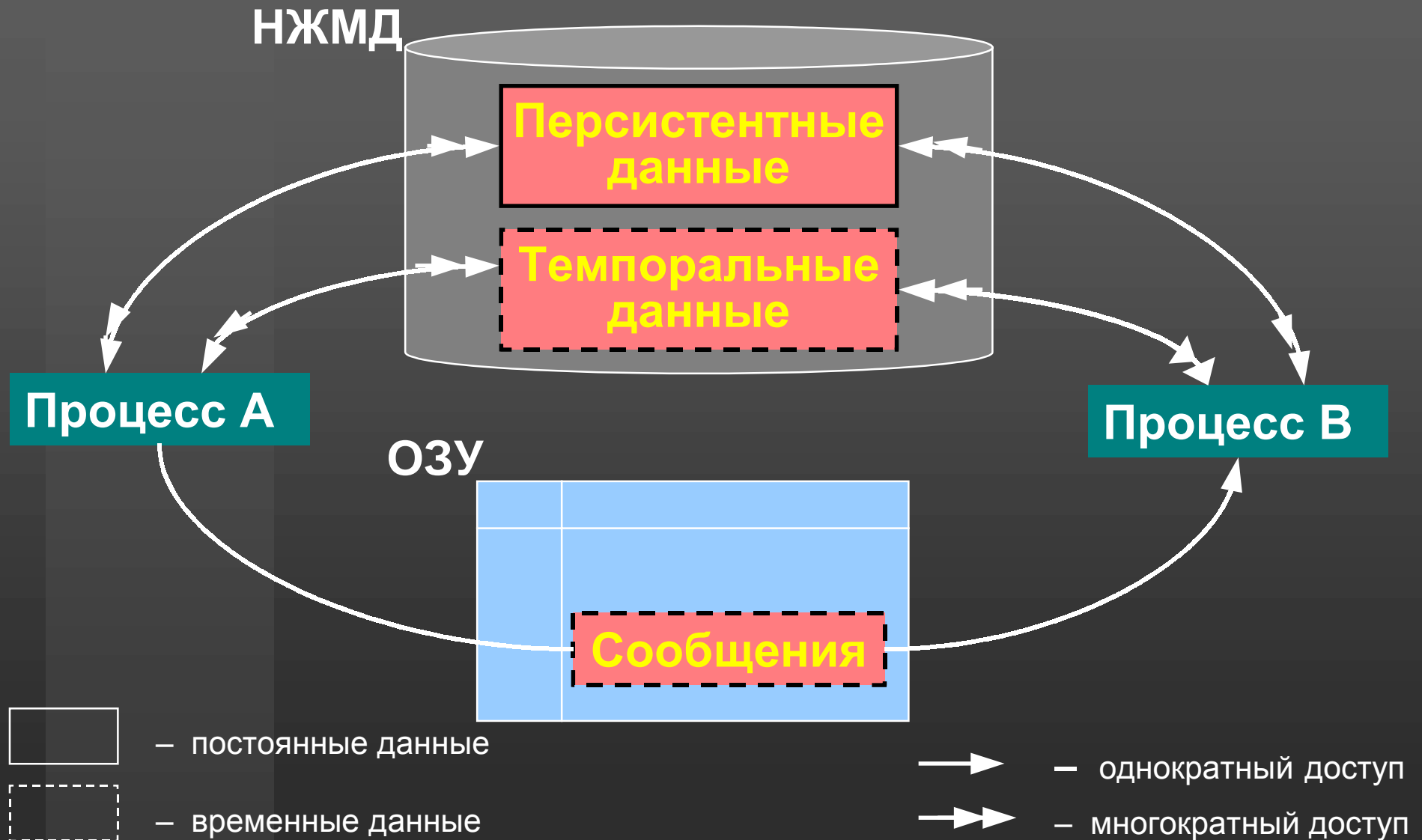
Примеры разбиения системы на кластеры



Характеристики данных

- Объем
- Время жизни
- Повторность использования

Классификация данных



Требования к системе управления данными

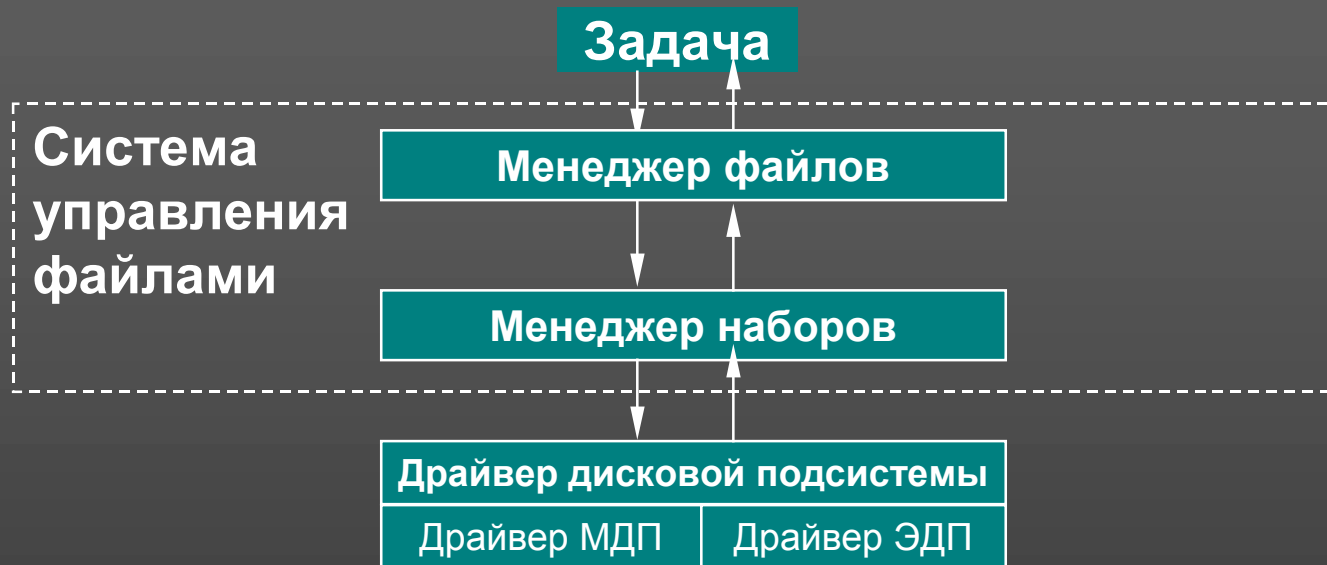
- Поддержка **логического уровня** структуризации вычислительной системы
- Обеспечение функций для эффективного управления **персистентными** данными
- Обеспечение функций для эффективного управления **темпоральными** данными
- Обеспечение функций для эффективного обмена данными **внутри процессорного кластера**
- Обеспечение функций для эффективного обмена данными **между процессорными кластерами**
- Поддержка **легковесных процессов**
- Поддержка **многопортовых обменов** данными

Структура комплекса Омега

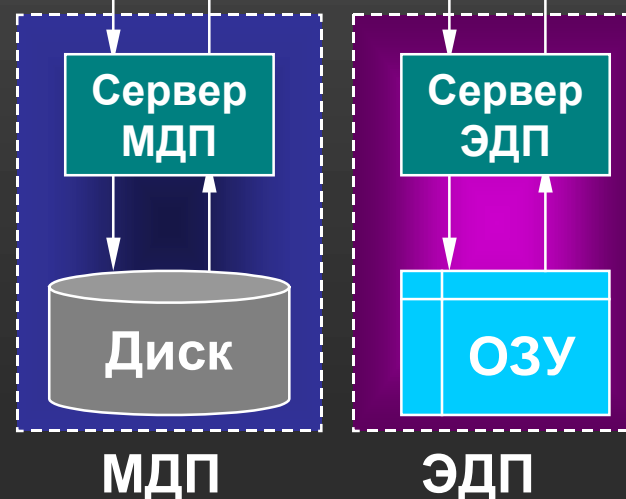


Система хранения данных

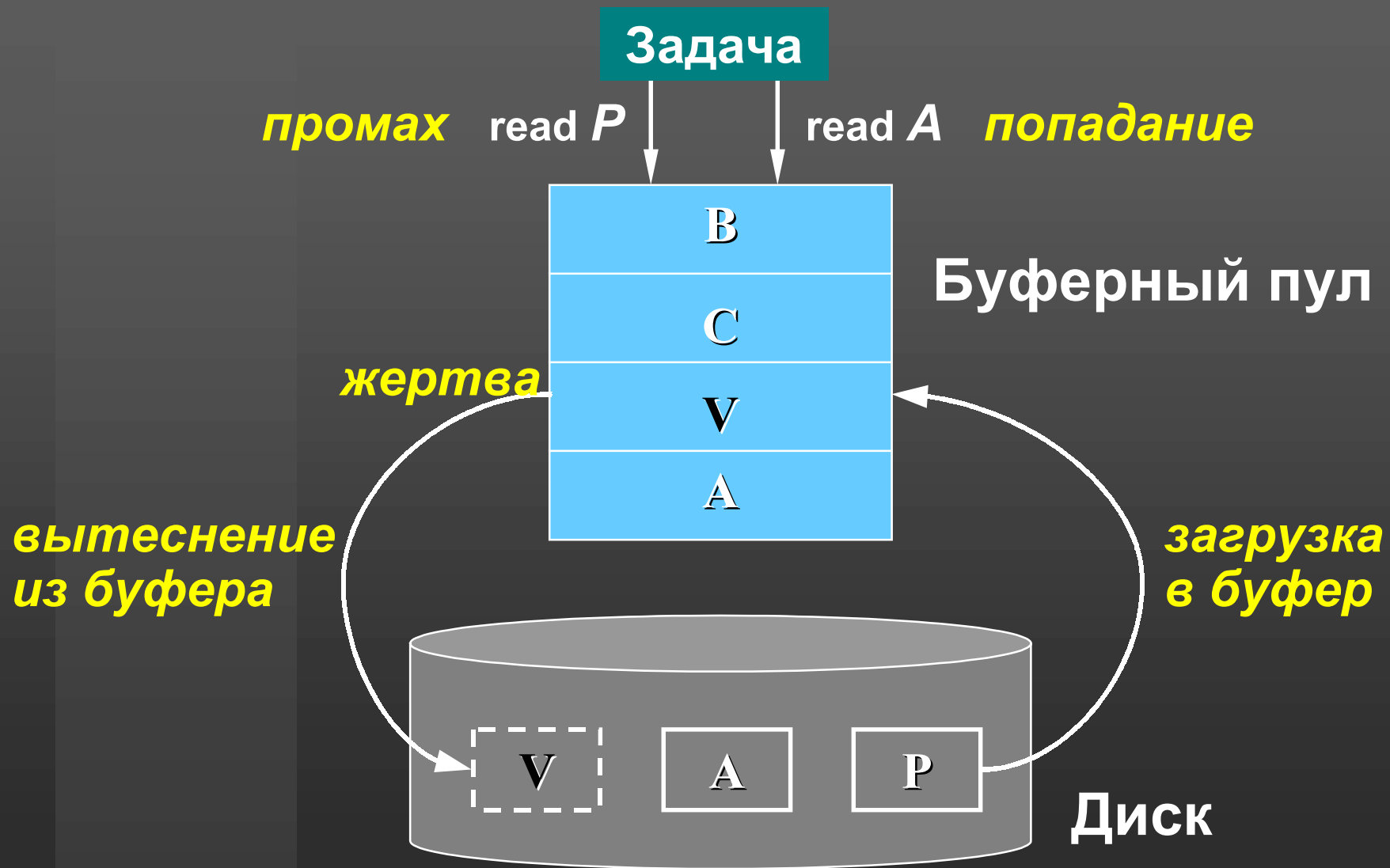
Узел-клиент



Узел-сервер



Буферизация данных



Общие стратегии вытеснения

Стратегия	Страница-жертва
OPT	К которой дольше всех не будет обращений
RANDOM	Случайная
LRU (Least Recently Used)	К которой дольше всего не было обращений
LFU (Least Frequently Used)	С наименьшим числом обращений
FIFO (First In First Out)	Первой помещенная в буфер
LIFO (Last In First Out)	Последней помещенная в буфер

- Не поддерживают избирательное вытеснение страниц
- Не адекватны в отдельных случаях доступа к данным

DIR-метод управления буферным пулом

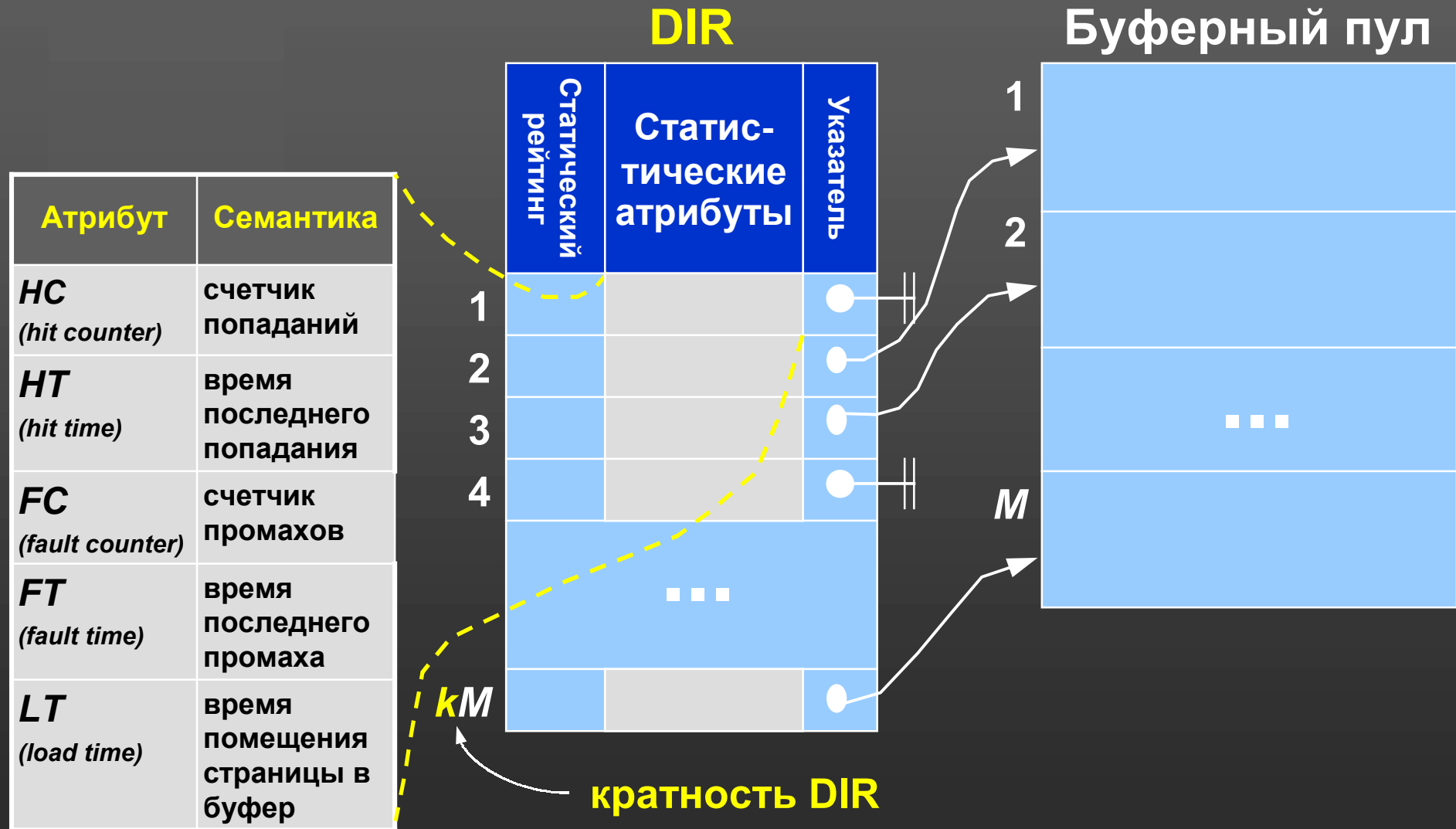
Требования к механизму вытеснения

- Поддержка избирательного вытеснения страниц
- Возможность динамического выбора стратегии вытеснения

Реализация

- Использование избыточного индекса буферного пула
- Введение рейтингов для образов страниц

Избыточный индекс буферного пула (DIR)



Рейтинг страницы

Рейтинг = статический + динамический

Жертва – страница с минимальным рейтингом

Статический рейтинг

- Реализует избирательное вытеснение страниц
- Целое число из $[0;20]$
- Атрибут открытого файла

Динамический рейтинг

- Реализует динамический выбор стратегии вытеснения страниц
- Функция $f(a_1, \dots, a_n) \rightarrow [0;1[$
- Атрибут страницы буфера

Моделирование стратегий вытеснения с помощью динамического рейтинга

Стратегия	Динамический рейтинг	Страница-жертва
LRU	$NORM(HT)$	К которой дольше всего не было обращений
LFU	$NORM(HC)$	С наименьшим числом обращений
FIFO	$NORM(LT)$	Первой помещенная в буфер
LIFO	$1/NORM(LT)$	Последней помещенная в буфер

- HT – время последнего попадания
- HC – счетчик попаданий
- LT – время помещения страницы в буфер
- $NORM$ – функция нормирования, приводящая значение к $[0;1[$

Примеры DIR-стратегий вытеснения

Стратегия	Динамический рейтинг
DIR_{FT}	$NORM(HT+FT/k)$
DIR_{FC}	$NORM(HC+FC/k)$
DIR_{LT}	$NORM(LT+FT/k)$
DIR_{HCFT}	$NORM(HC+FT)$

- $NORM$ – функция нормирования, приводящая значение к $[0;1[$
- k – кратность DIR

Численные эксперименты

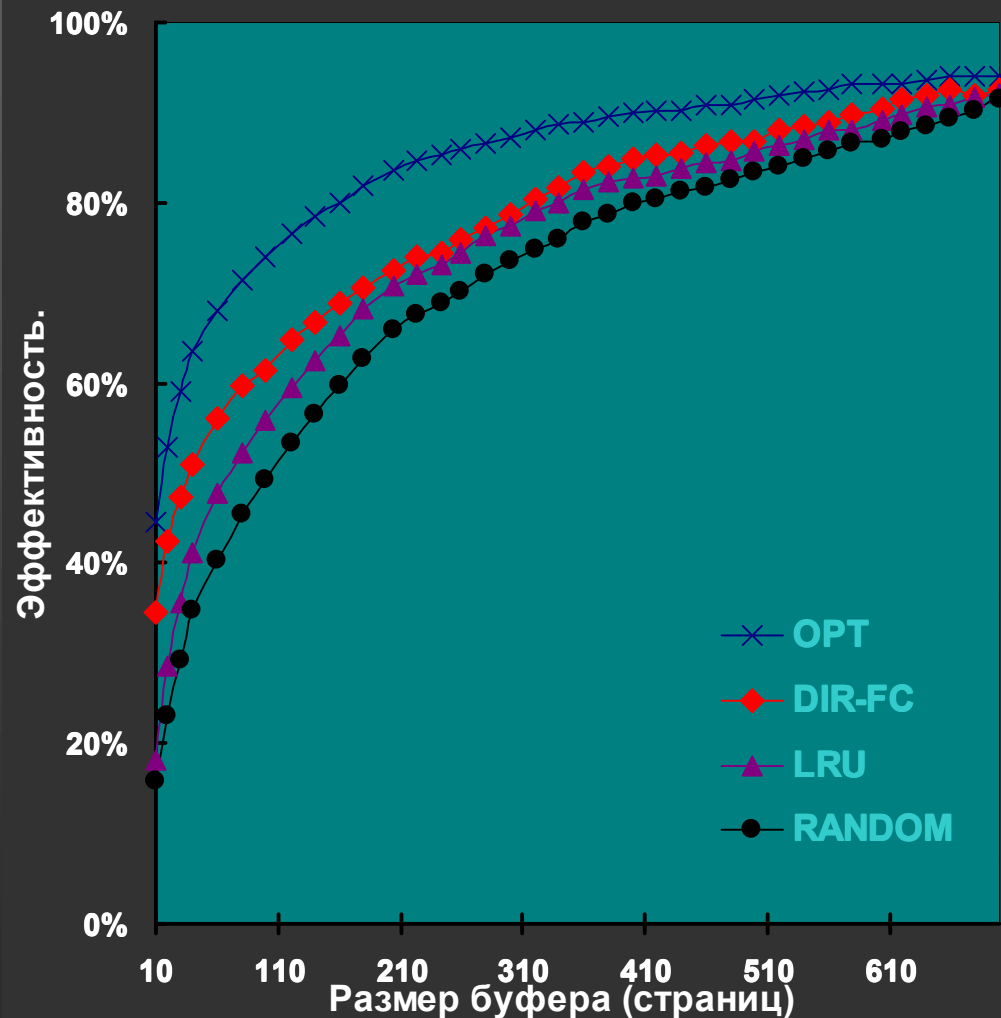
Цели

- Сравнение эффективности общих стратегий с DIR-стратегией
- Сравнение эффективности различных DIR-стратегий
- Исследование влияния кратности DIR на эффективность DIR-стратегии

Параметры

Параметр	Семантика	Значение
L	Количество обращений к страницам диска	3000
N	Количество страниц диска, к которым осуществляются обращения	1000
$Z(i)$	Функция распределения частот обращений	$\left(\frac{i}{N}\right)^{\log 0,8 / \log 0,2}$
k	Кратность DIR	10

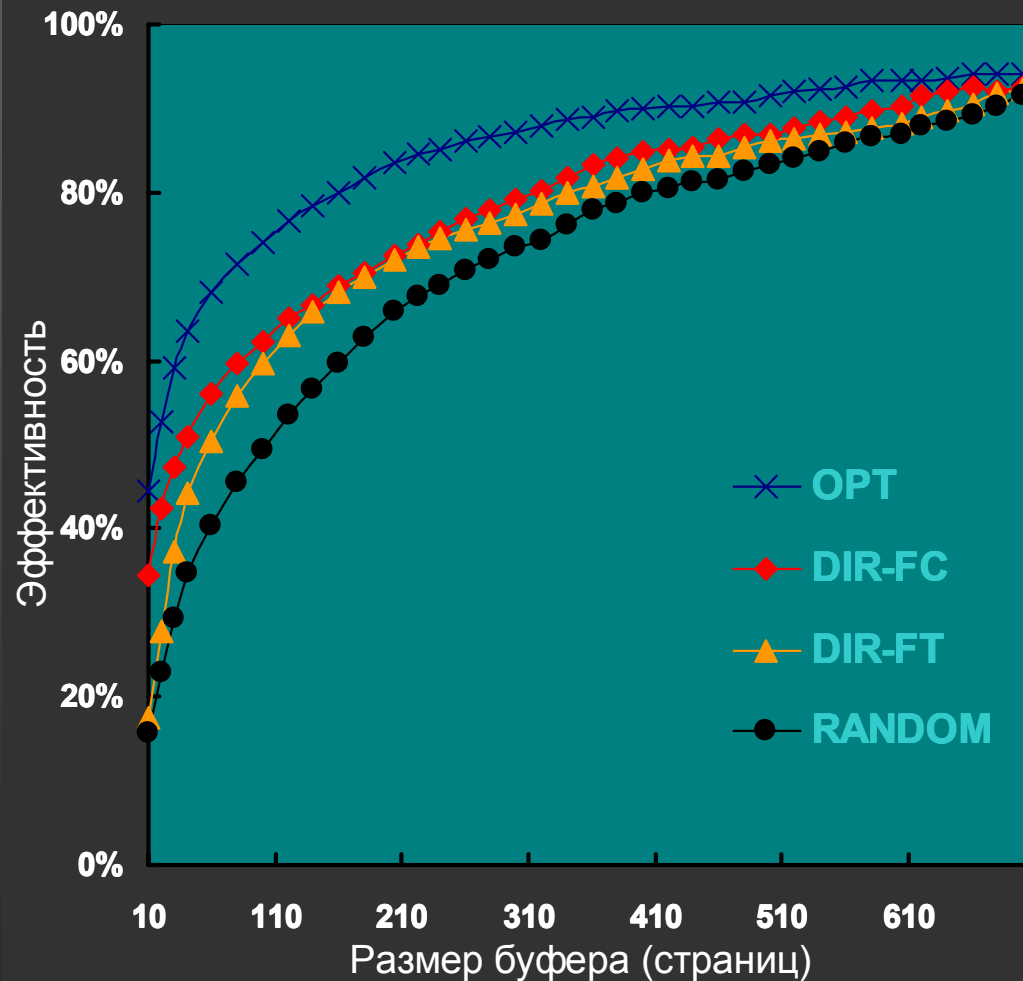
Сравнение эффективности общих стратегий с DIR-стратегией



Стратегия	Динамический рейтинг
DIR_{FC}	$NORM(HC+FC/k)$
LRU	$NORM(HT)$

Показана **более высокая эффективность DIR-стратегии** по сравнению с классической стратегией LRU

Сравнение эффективности различных DIR-стратегий



Стратегия	Динамический рейтинг
DIR_{FC}	$NORM(HC+FC/k)$
DIR_{FT}	$NORM(HT+FT/k)$

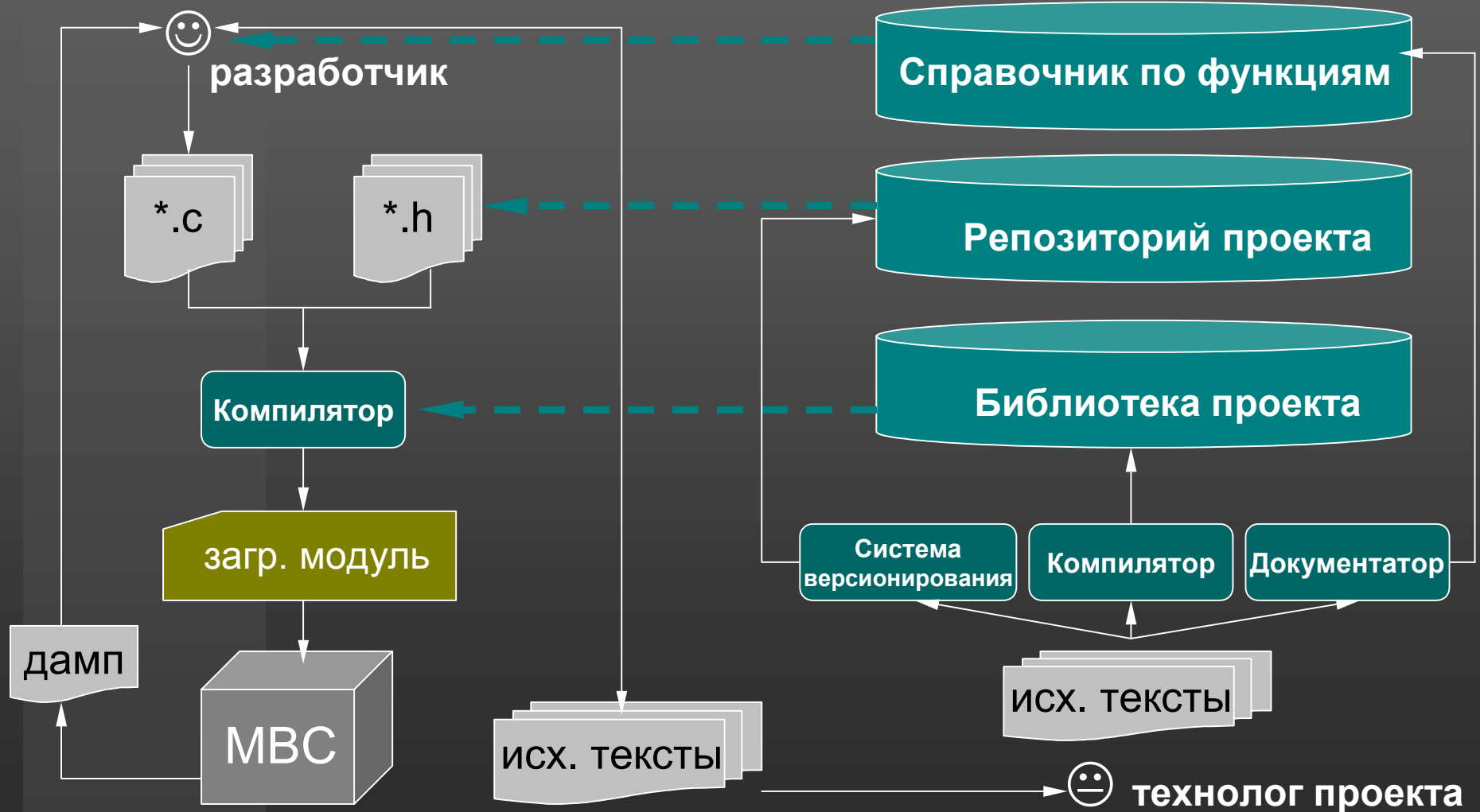
Показана **более высокая эффективность** стратегии DIR_{FC}

Исследование влияния кратности DIR на эффективность DIR-стратегии

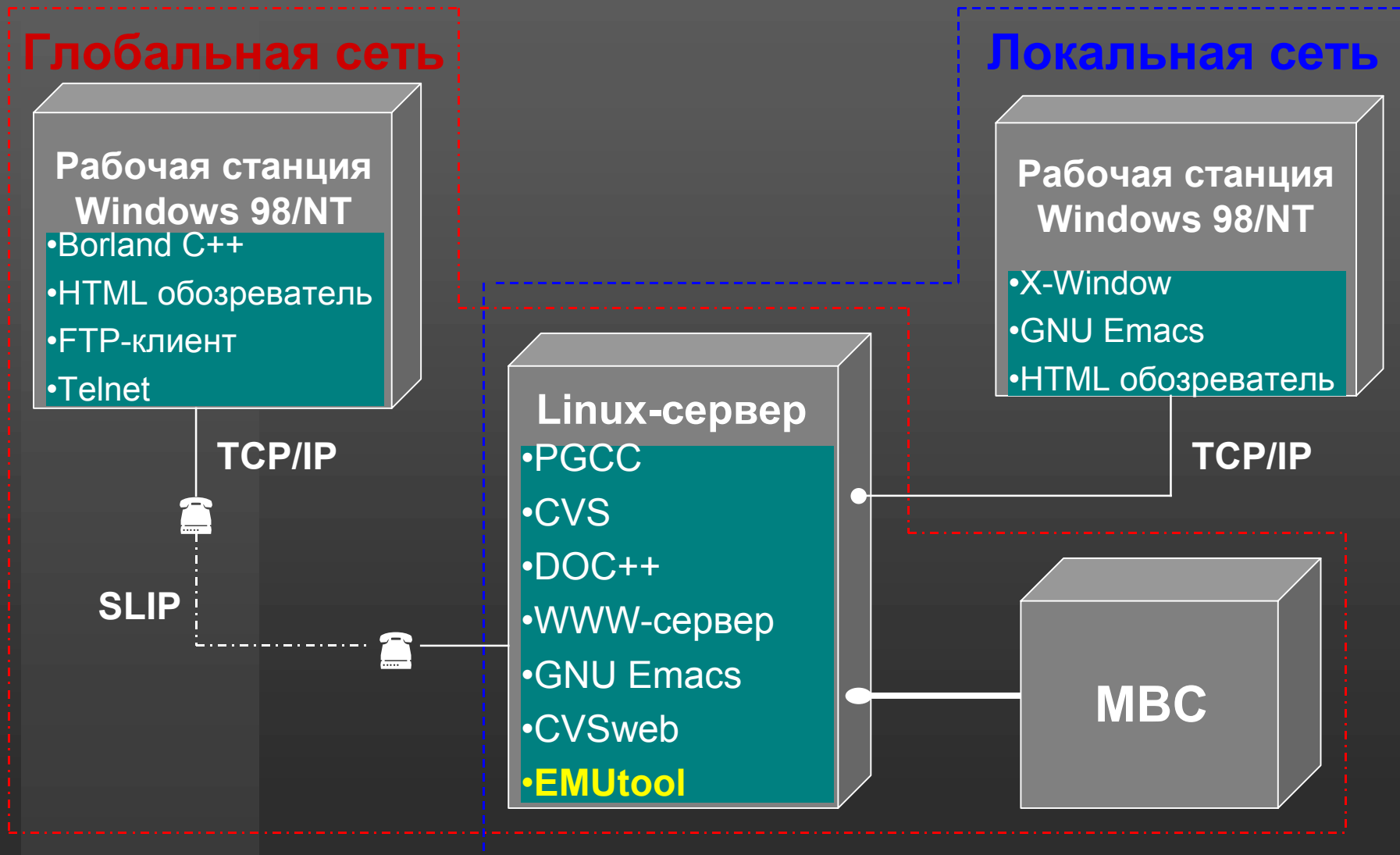
Размер буфера Эффективность DIR_{FC} , %
k=5 k=10 **k=20** k=25 k=30

10	33,70	34,50	34,11	34,40	34,45
20	42,20	42,40	41,95	40,40	41,80
30	46,90	47,40	46,55	44,90	46,55
40	49,60	51,00	51,30	49,40	49,90
60	54,30	56,00	59,40	55,60	54,41
80	57,40	59,70	63,50	59,80	59,00
100	60,40	62,10	66,90	63,80	62,60
120	62,80	64,80	69,80	66,90	65,59
140	64,70	66,70	72,20	69,80	68,60
160	66,50	68,90	73,60	71,40	70,50
180	68,70	70,50	74,80	73,40	72,28
200	70,80	72,60	76,20	74,60	74,00
220	71,60	73,90	77,30	76,00	75,40
240	72,90	74,70	79,00	77,20	77,20
260	73,80	76,00	80,00	78,20	77,80
280	75,70	77,10	80,90	79,00	79,00
300	76,55	78,70	81,80	81,10	81,10
320	78,30	80,30	82,00	81,90	82,00
340	79,80	81,70	82,70	82,20	82,10
360	80,80	82,30	83,70	82,50	82,40
380	81,70	83,10	84,30	82,70	82,70
400	82,80	84,00	84,60	82,90	82,90
420	83,60	84,40	84,80	83,20	83,10
440	84,80	85,30	85,00	83,90	83,50
460	85,00	86,40	85,40	84,10	84,00
480	85,50	86,80	85,80	84,60	84,20
500	86,20	86,90	86,95	85,20	85,00

Технология коллективной разработки



Интегрированная среда разработки



Основные результаты

- Предложена **методика построения программного комплекса для управления данными** в вычислительных системах с массовым параллелизмом.
- Разработана **система хранения данных** для многопроцессорного вычислительного комплекса МВС-100.
- Разработан **метод управления буферным пулом (DIR-метод)**, базирующийся на введении рейтингов страниц и использовании избыточного индекса буферного пула.
- Предложена **методика построения эффективных стратегий вытеснения** страниц, базирующихся на DIR-методе. Проведены **численные эксперименты**, подтверждающие более высокую эффективность DIR-стратегий по сравнению с классическими.
- Предложена **технология разработки больших программных комплексов** для многопроцессорной вычислительной системы МВС-100.
- Система хранения данных **внедрена в опытную эксплуатацию** в ряде российских организаций.