

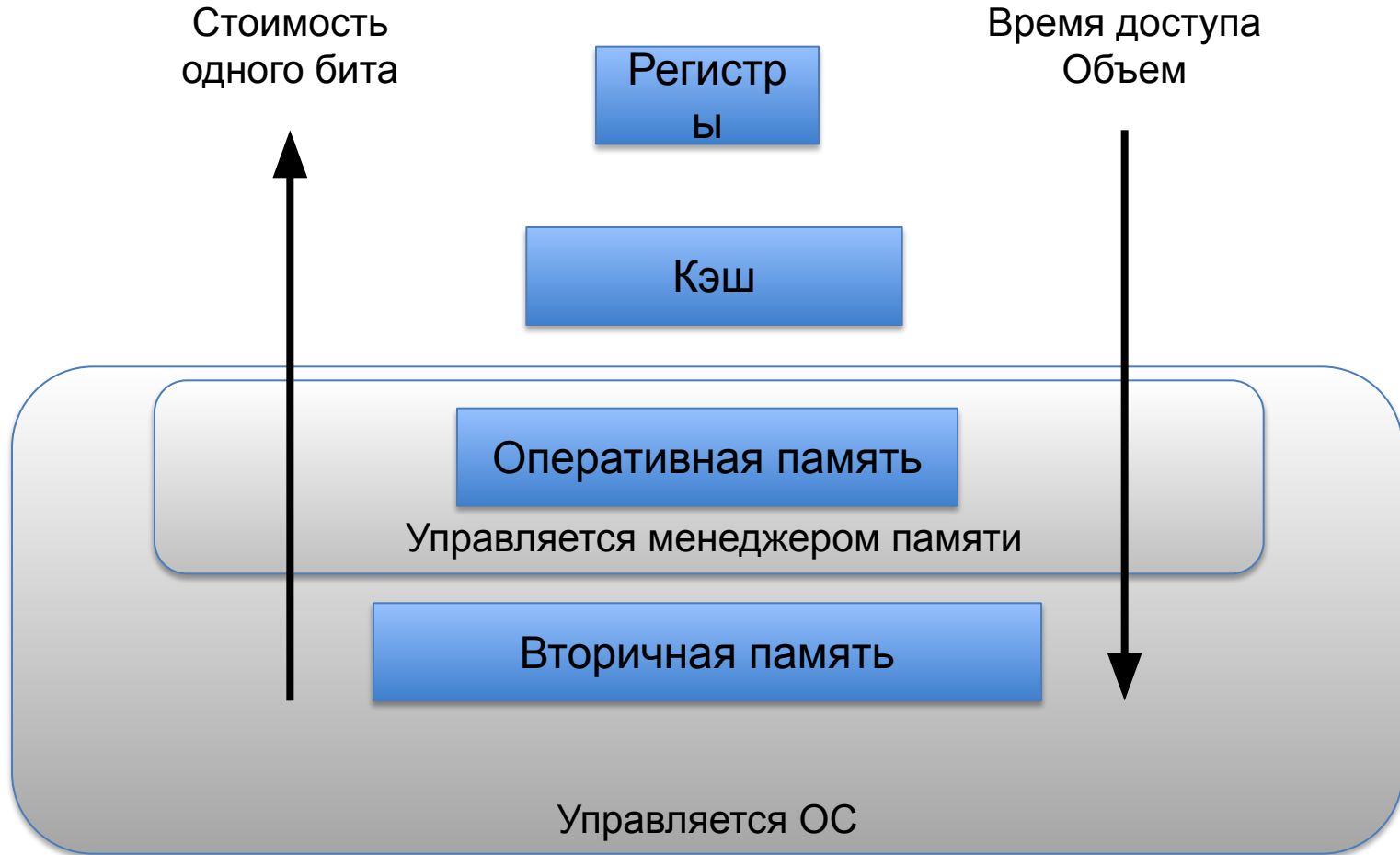
Управление памятью

Простые схемы управления
памятью

Лекция 12,13
2020

Управление памятью

Иерархия памяти



Управление памятью

Принцип локальности

Большинство реальных программ в течение некоторого отрезка времени работает с небольшим набором адресов памяти – это *принцип локальности*

Принцип локальности связан с особенностями человеческого мышления

Управление памятью

Проблема разрешения адресов

Человеку свойственно символическое мышление. Адреса (имена) переменных описываются идентификаторами, формируя символическое адресное пространство

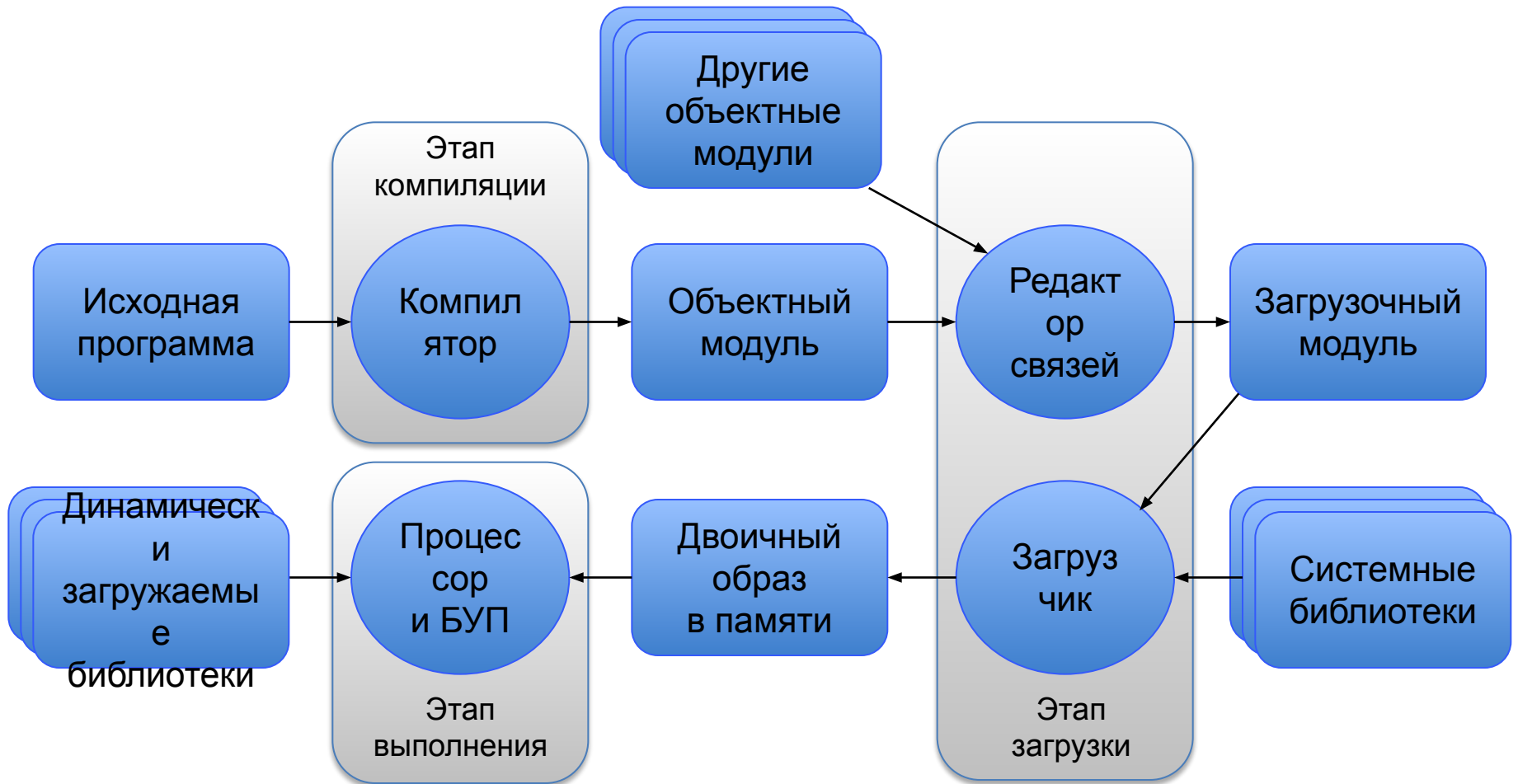
Как ? ↓ Когда ?

Оперативная физическая память может быть представлена в виде массива ячеек с линейными адресами

Совокупность всех доступных физических адресов в вычислительной системе – это ее физическое адресное пространство

Управление памятью

Связывание адресов



Управление памятью

Логическое адресное пространство

Символьное адресное пространство – совокупность всех допустимых идентификаторов переменных



Логическое адресное пространство – совокупность всех допустимых адресов, с которыми работает процессор



Физическое адресное пространство – совокупность всех доступных физических адресов в вычислительной системе

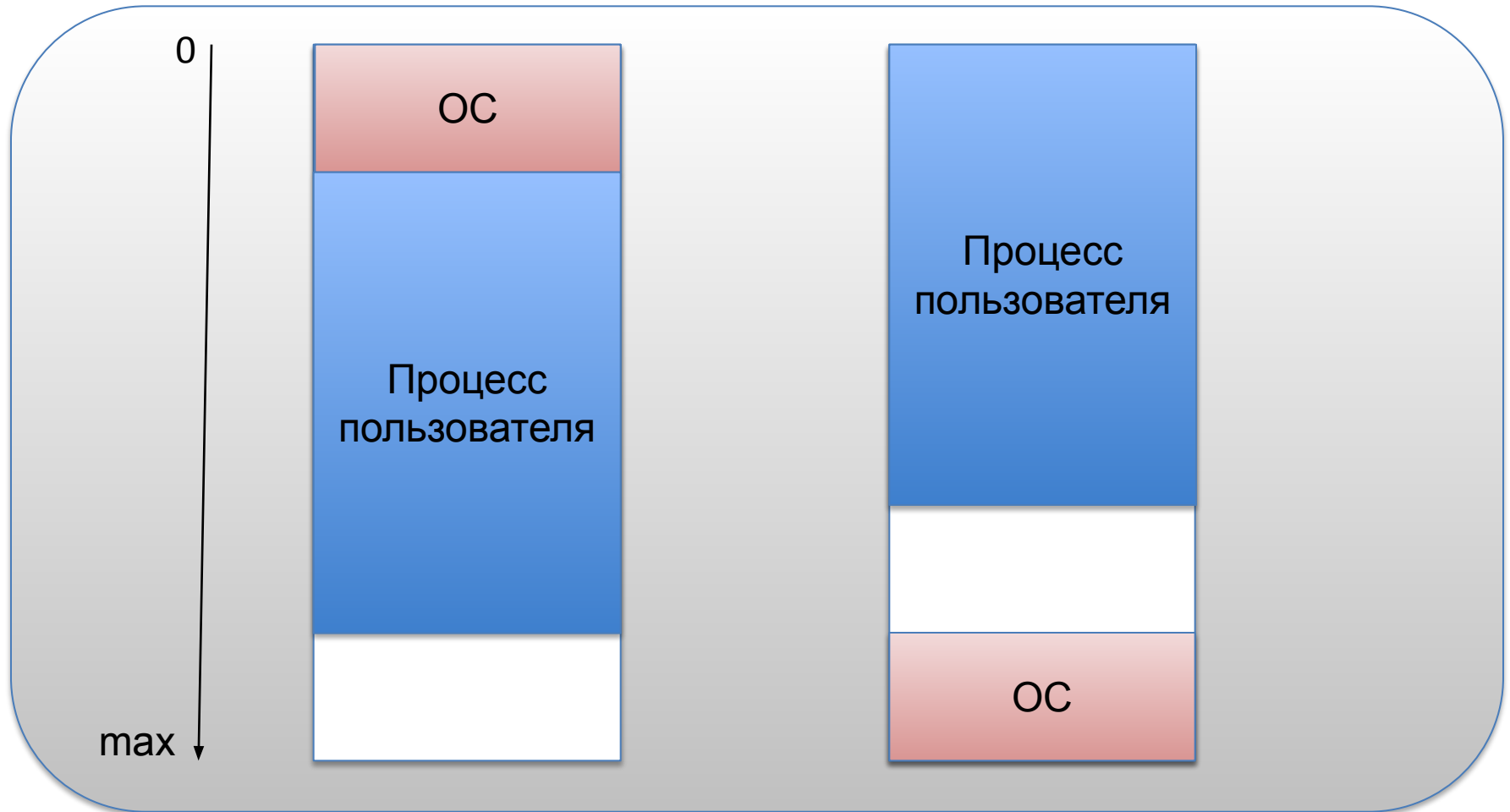
Управление памятью

Функции ОС и hardware

1. Отображение логического адресного пространства процесса на физическое адресное пространство
2. Распределение памяти между конкурирующими процессами
3. Контроль доступа к адресным пространствам процессов
4. Выгрузка процессов (целиком или частично) во внешнюю память
5. Учет свободной и занятой памяти

Простые схемы управления памятью

Однопрограммная система



Простые схемы управления памятью

Организация больших программ

- **Оверлейная структура**

Программа разбивается на несколько частей. Постоянно в памяти находится только загрузчик оверлеев, небольшое количество общих данных и процедур, а части загружаются по очереди

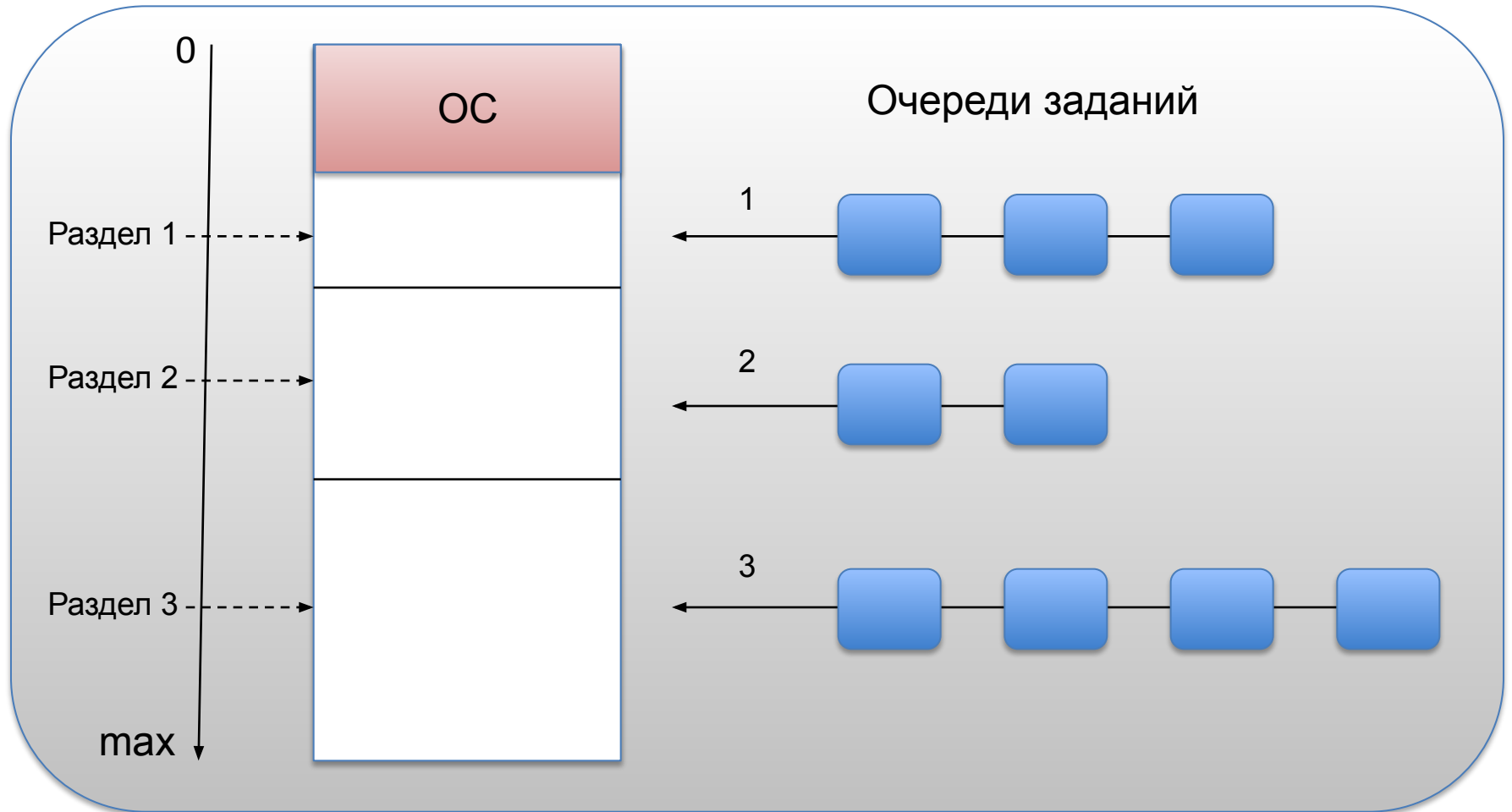
- **Динамическая загрузка процедур**

Процедуры загружаются в память только по мере необходимости, после обращения к ним

Оба способа основаны на применении принципа
локальности

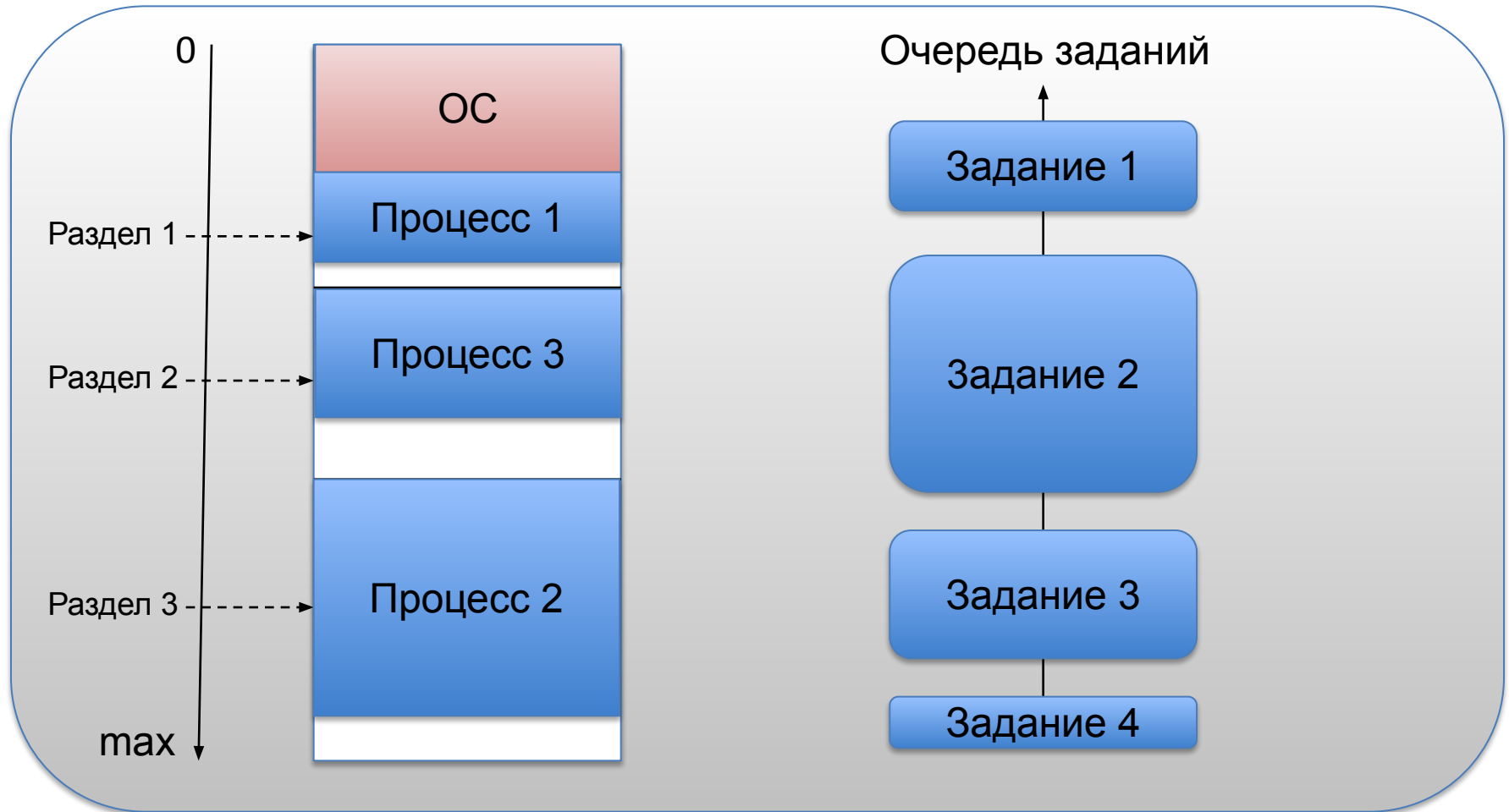
Простые схемы управления памятью

Фиксированные разделы



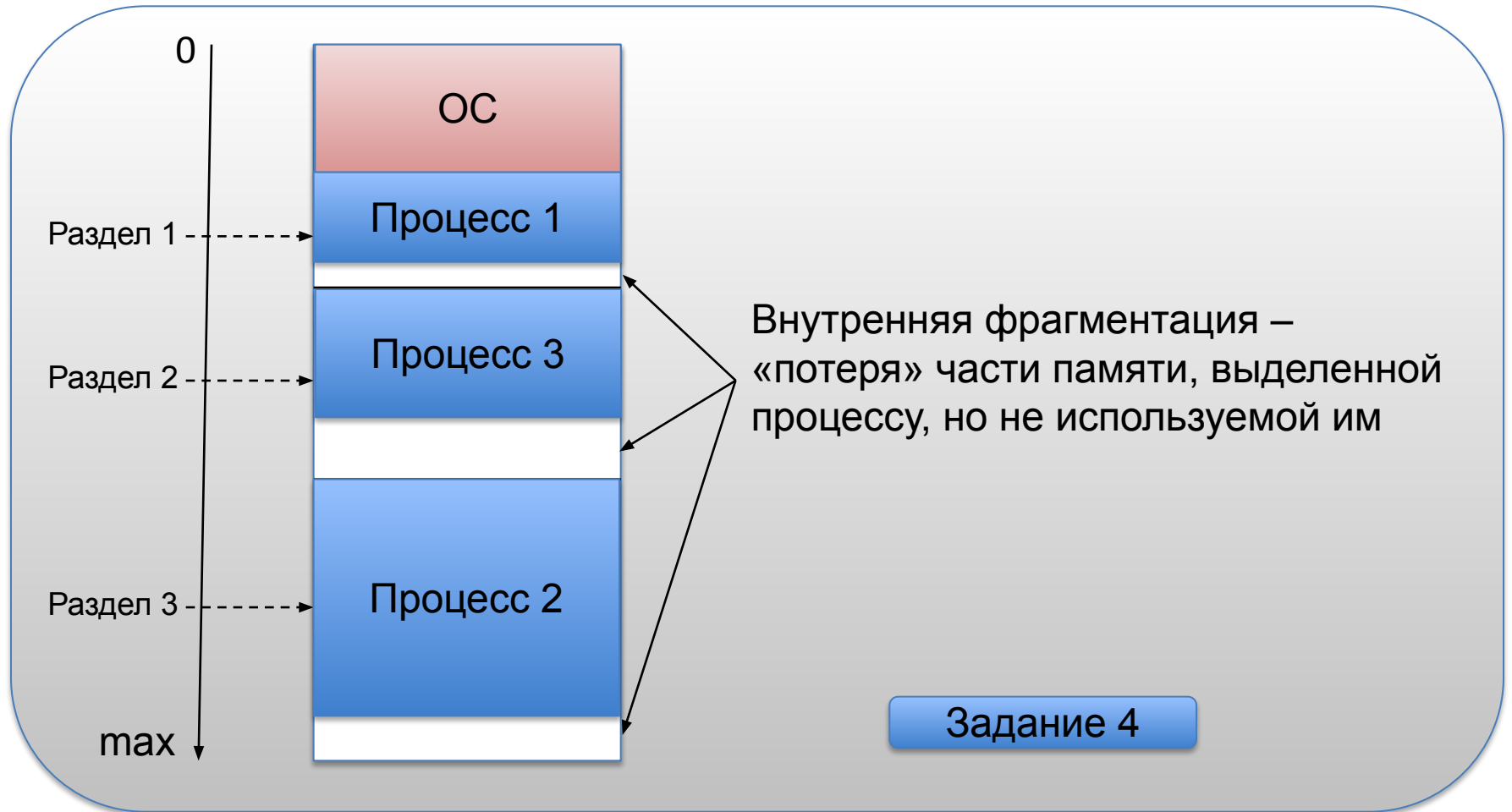
Простые схемы управления памятью

Фиксированные разделы



Простые схемы управления памятью

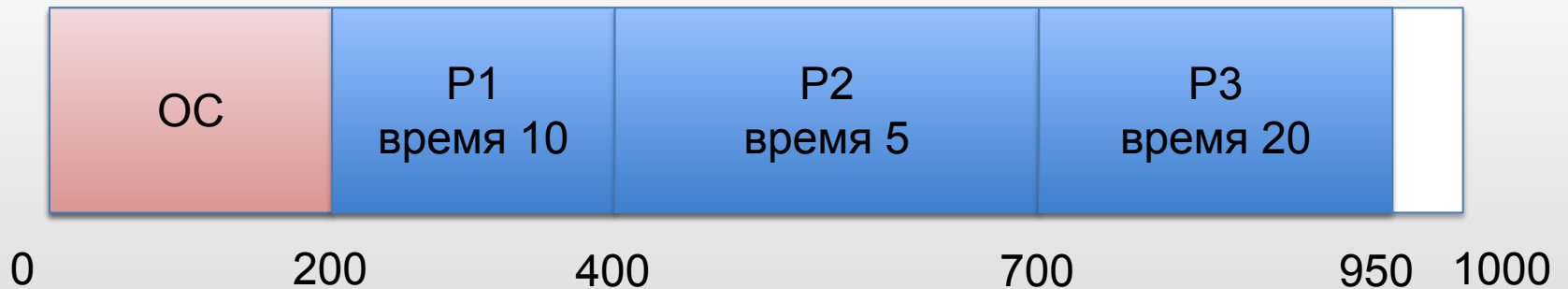
Фиксированные разделы



Простые схемы управления памятью

Динамические разделы

Память



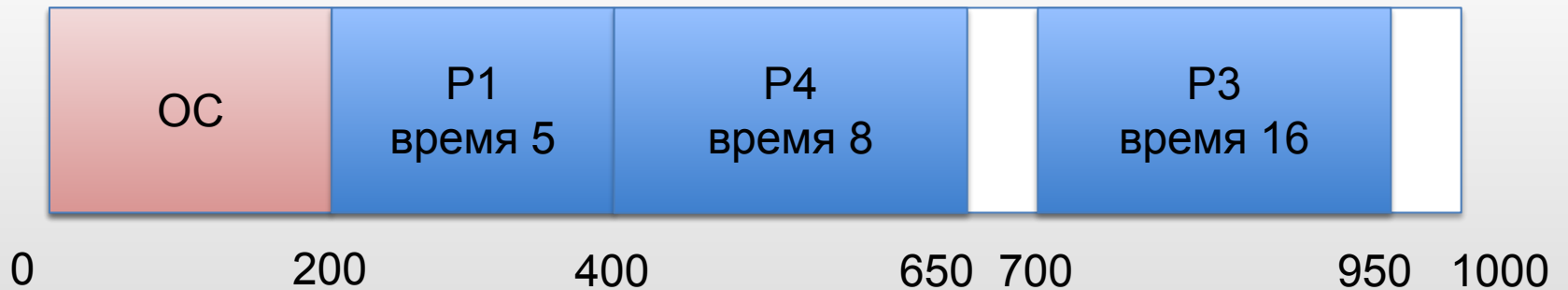
Очередь заданий

№	1	2	3	4	5
память	200	300	250	250	70
время	10	5	20	8	15

Простые схемы управления памятью

Динамические разделы

Память



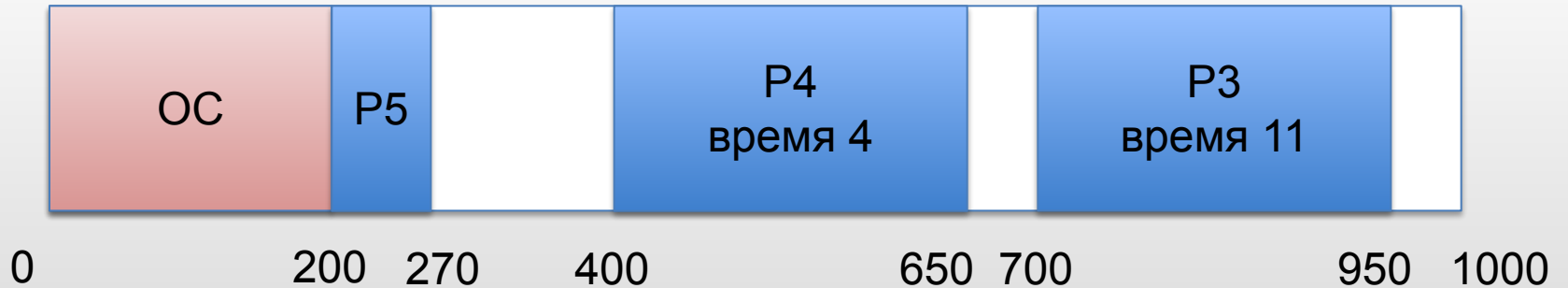
Очередь заданий

№				4	5
память				250	70
время				8	15

Простые схемы управления памятью

Динамические разделы

Память



Очередь заданий

№					5
память					70
время					15

Простые схемы управления памятью

Динамические разделы

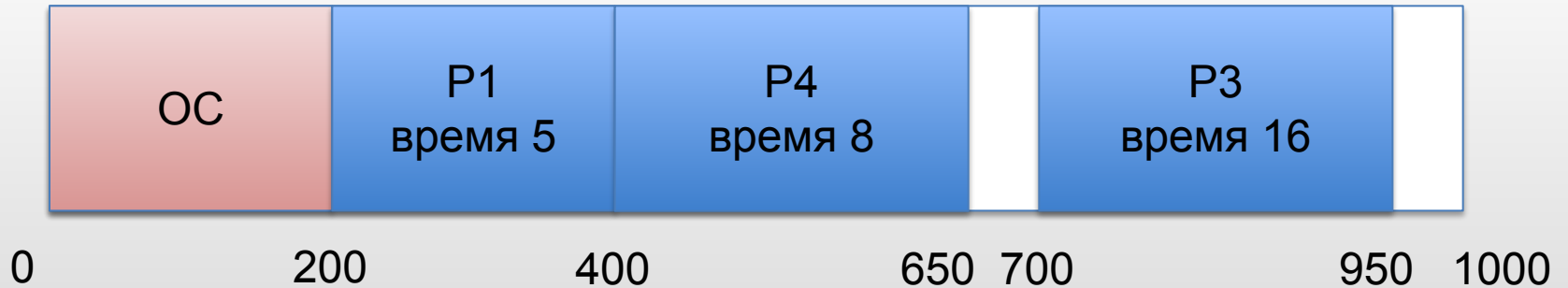
Стратегии размещения нового процесса в памяти

- Первый подходящий (first-fit). Процесс размещается в первое подходящее по размеру пустое место
- Наиболее подходящий (best-fit). Процесс размещается в наименьшее подходящее по размеру пустое место
- Наименее подходящий (worst-fit). Процесс размещается в наибольшее пустое место

Простые схемы управления памятью

Динамические разделы

Память



Очередь заданий

№					5
память					70
время					15

Простые схемы управления памятью

Динамические разделы

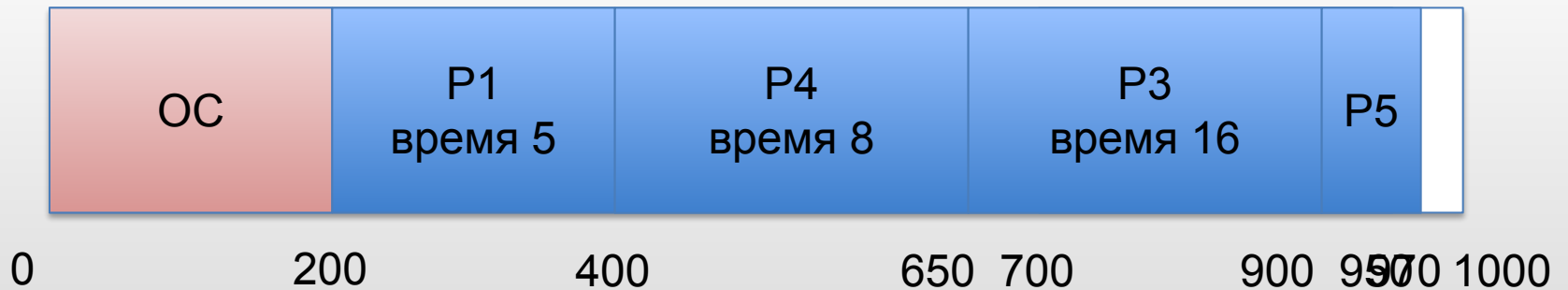


Внешняя фрагментация – невозможность использования памяти, неиспользуемой процессами, из-за ее раздробленности

Простые схемы управления памятью

Динамические разделы

Сборка мусора



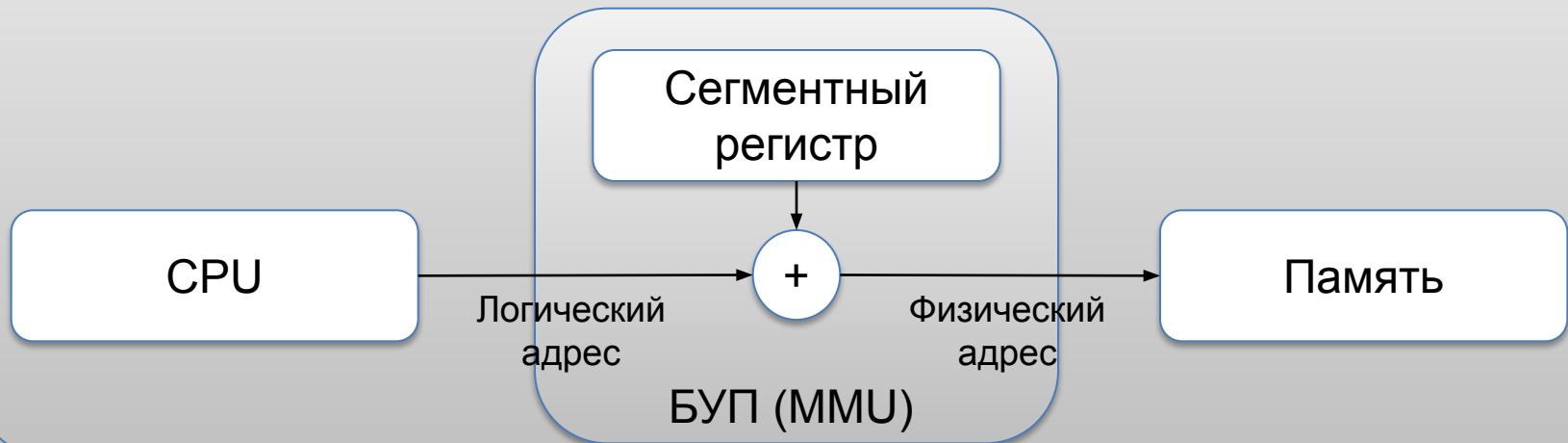
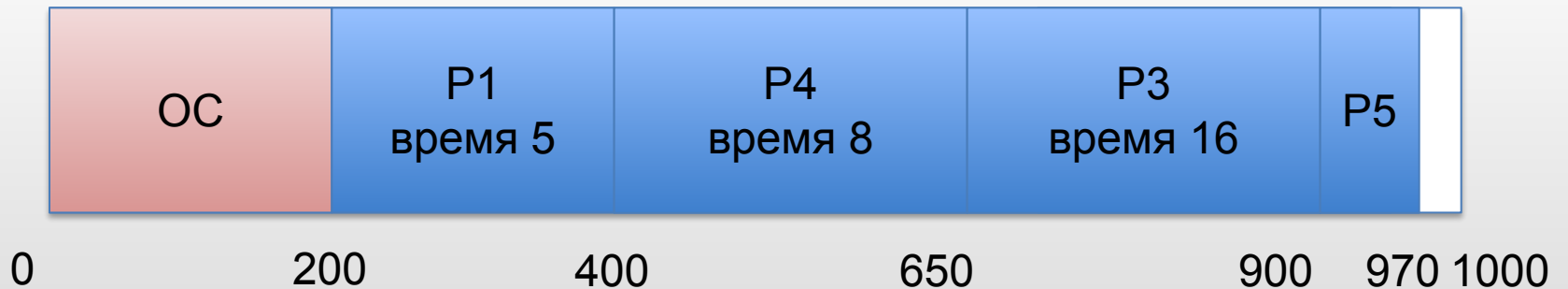
Очередь заданий

№					5
память					70
время					15

Простые схемы управления памятью

Динамические разделы

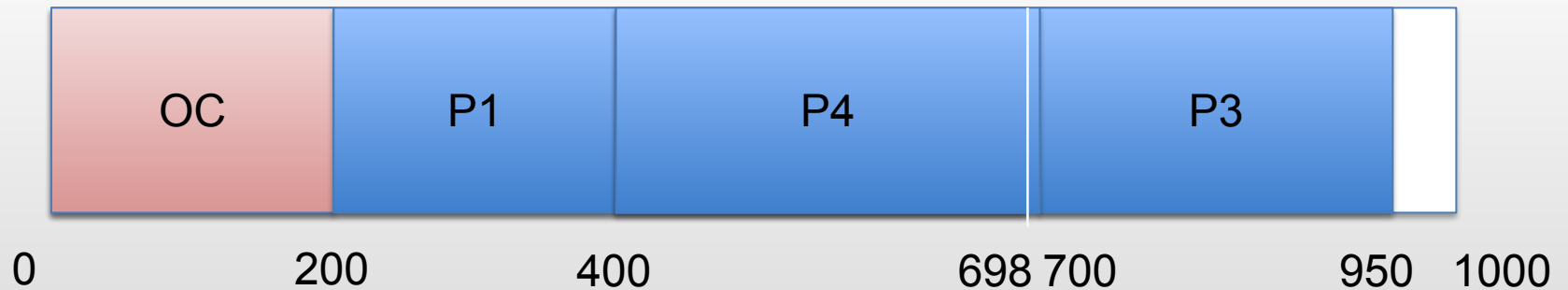
Сборка мусора



Простые схемы управления памятью

Динамические разделы

Память

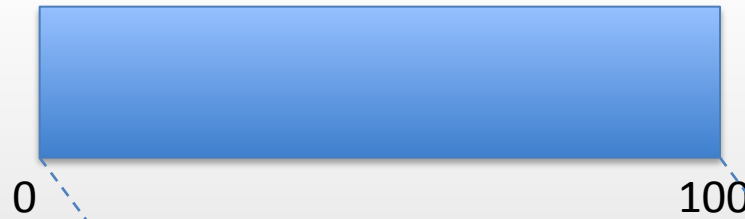


Возможна и внутренняя фрагментация при почти полном заполнении процессом пустого фрагмента

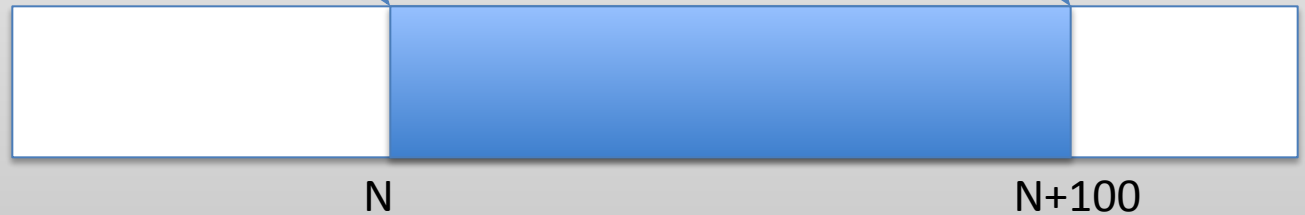
Простые схемы управления памятью

Линейное непрерывное отображение

Логическое
адресное
пространство

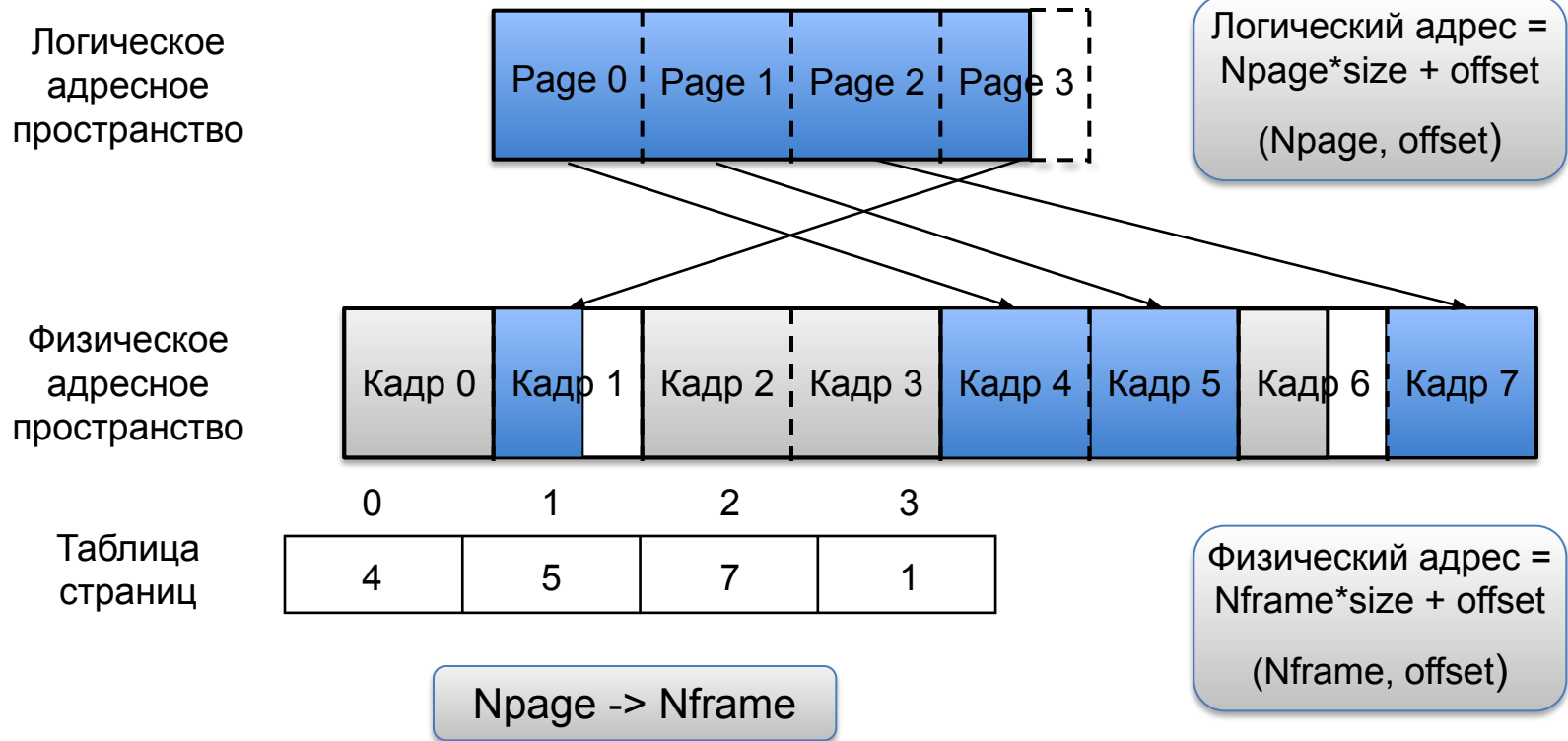


Физическое
адресное
пространство



Кусочно-непрерывное отображение

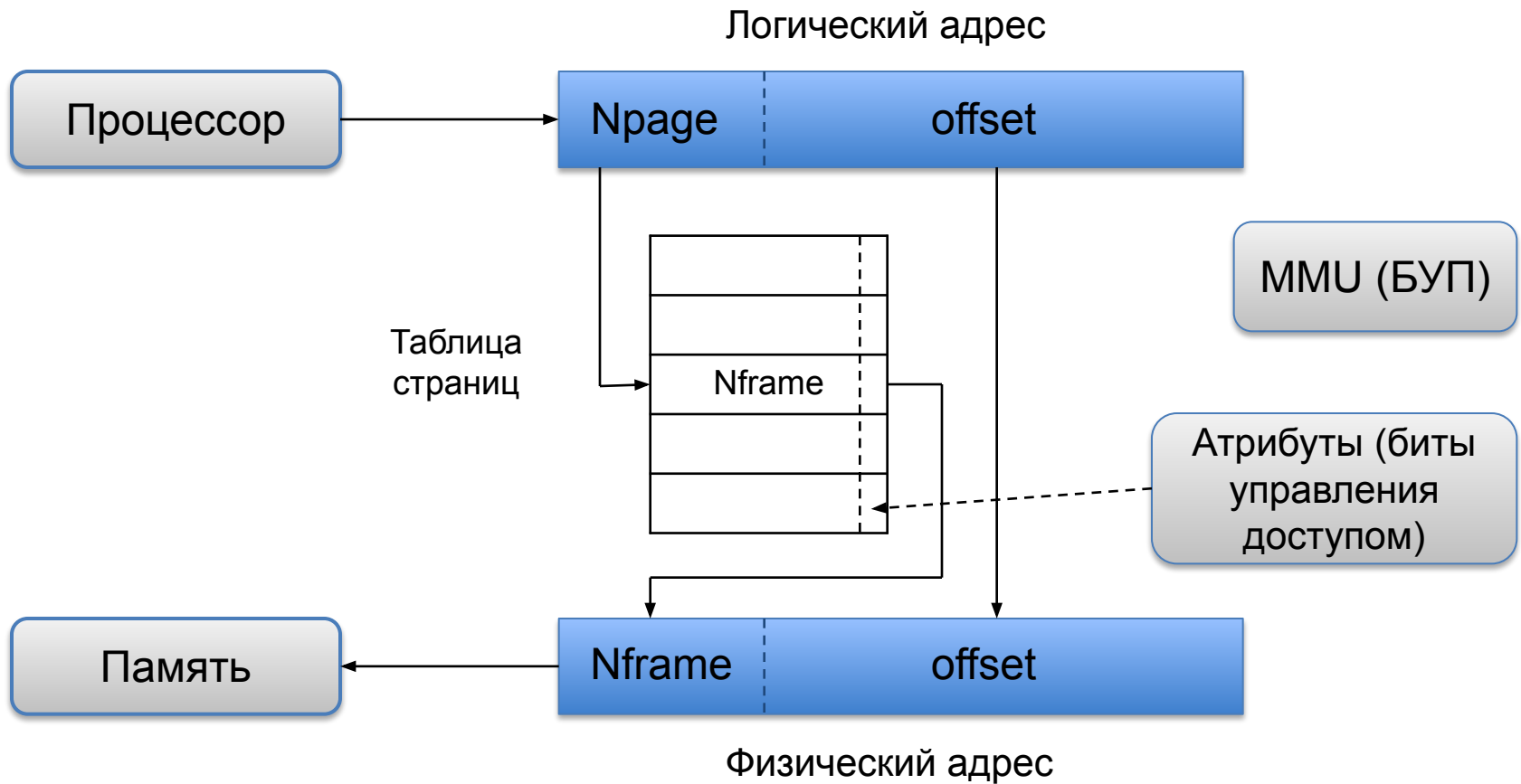
Страничная организация памяти



Свойственна внутренняя фрагментация

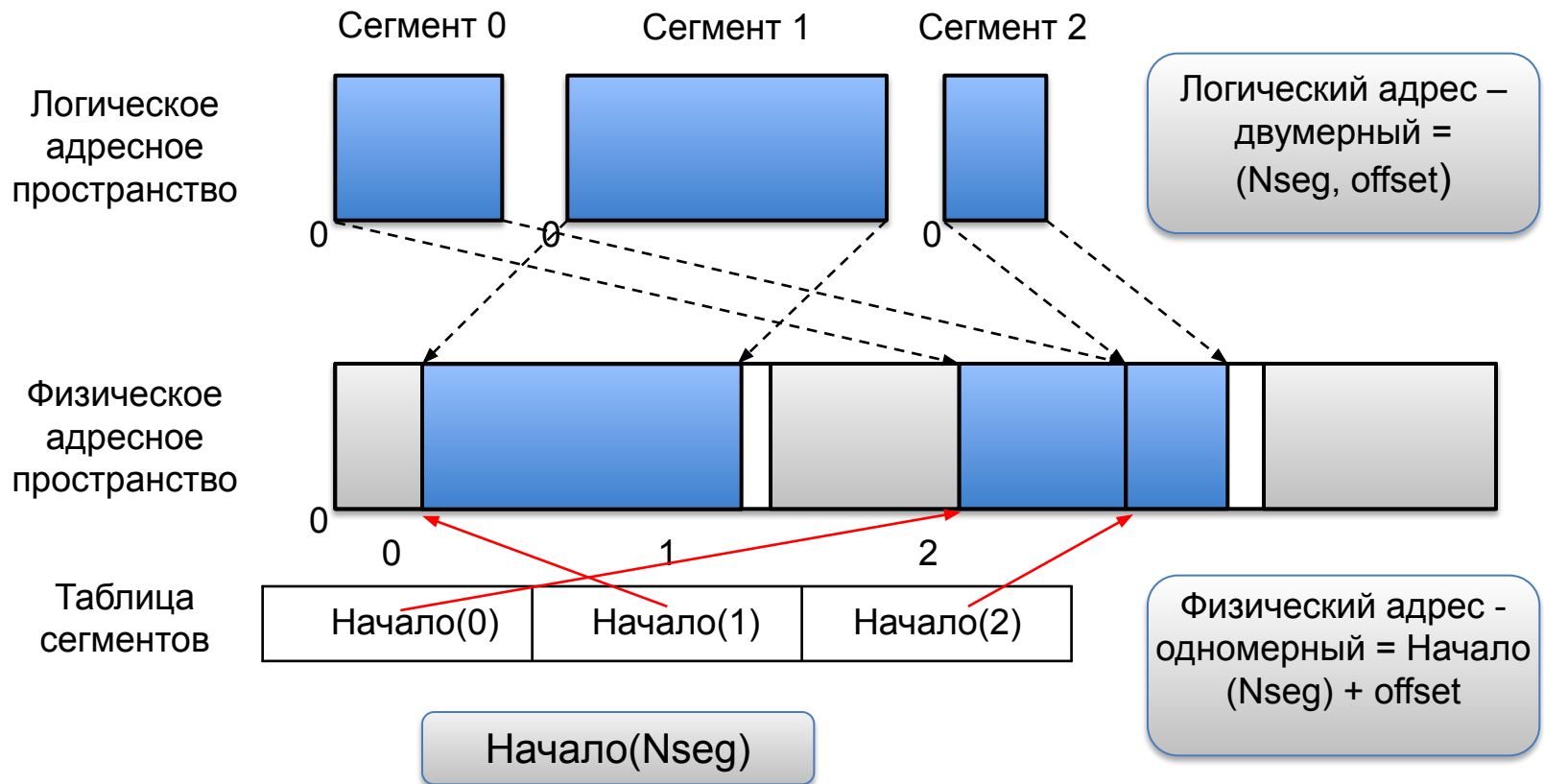
Кусочно-непрерывное отображение

Страничная организация памяти



Кусочно-непрерывное отображение

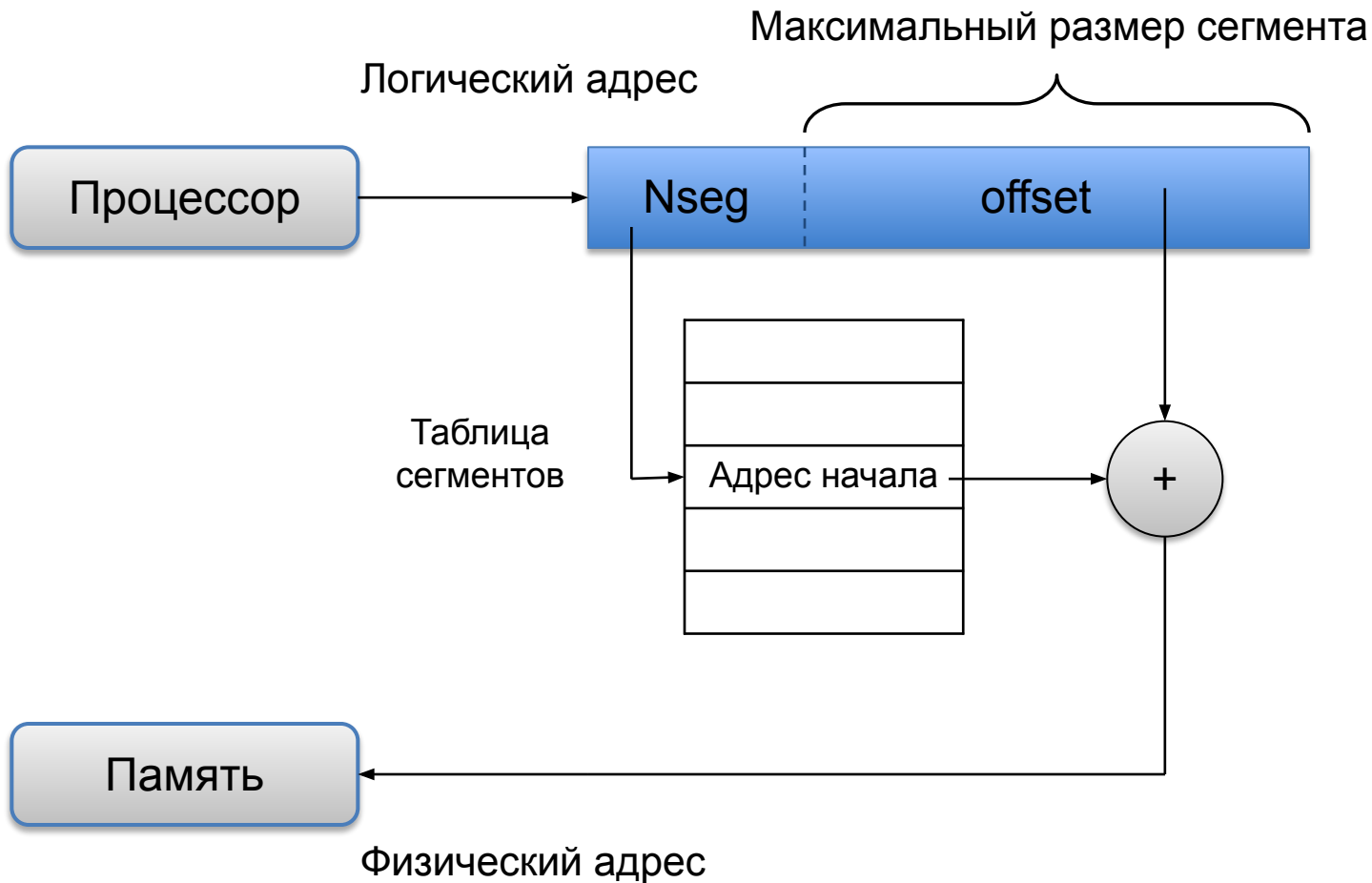
Сегментная организация памяти



Свойственна внешняя фрагментация

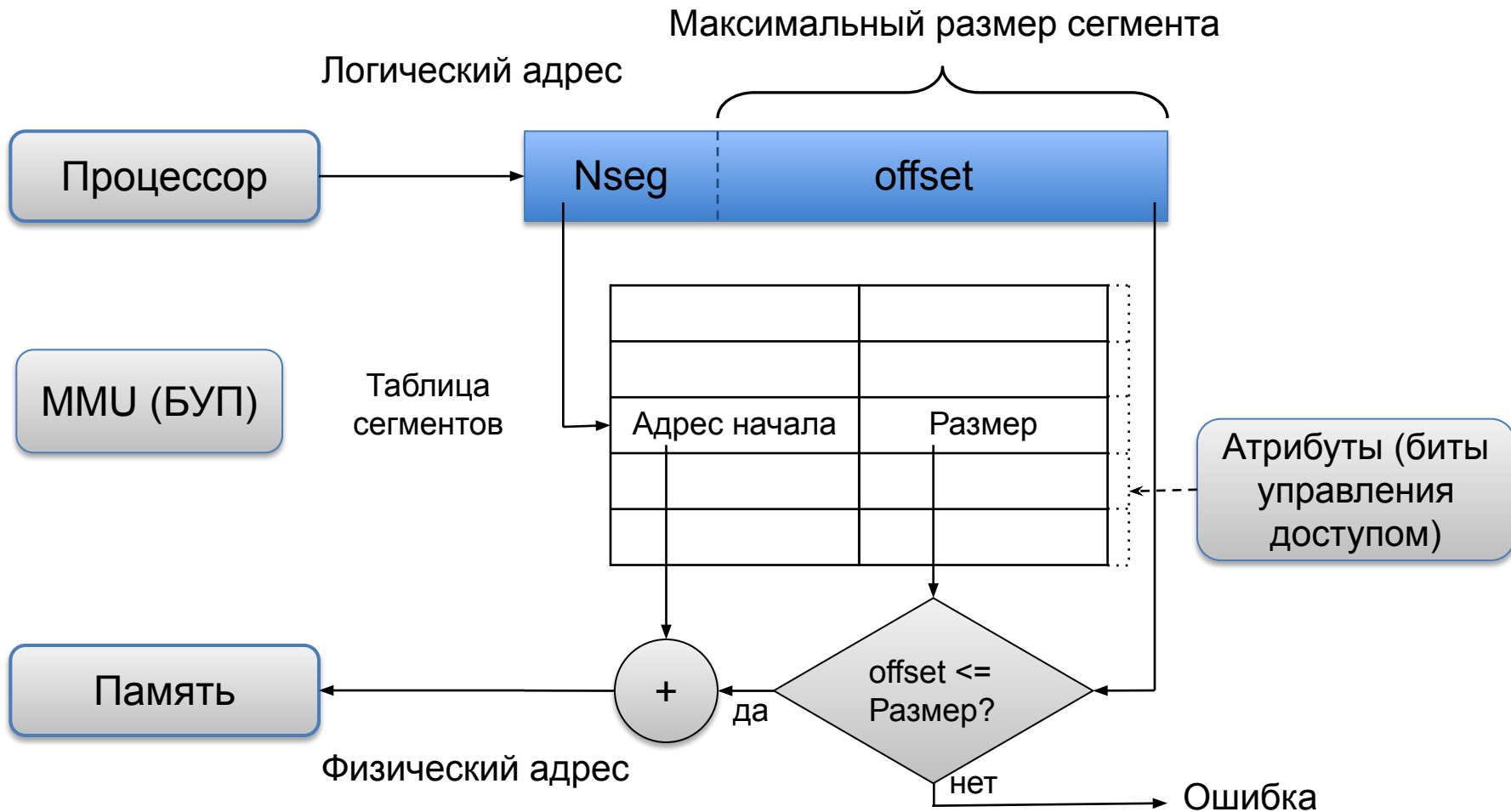
Кусочно-непрерывное отображение

Сегментная организация памяти



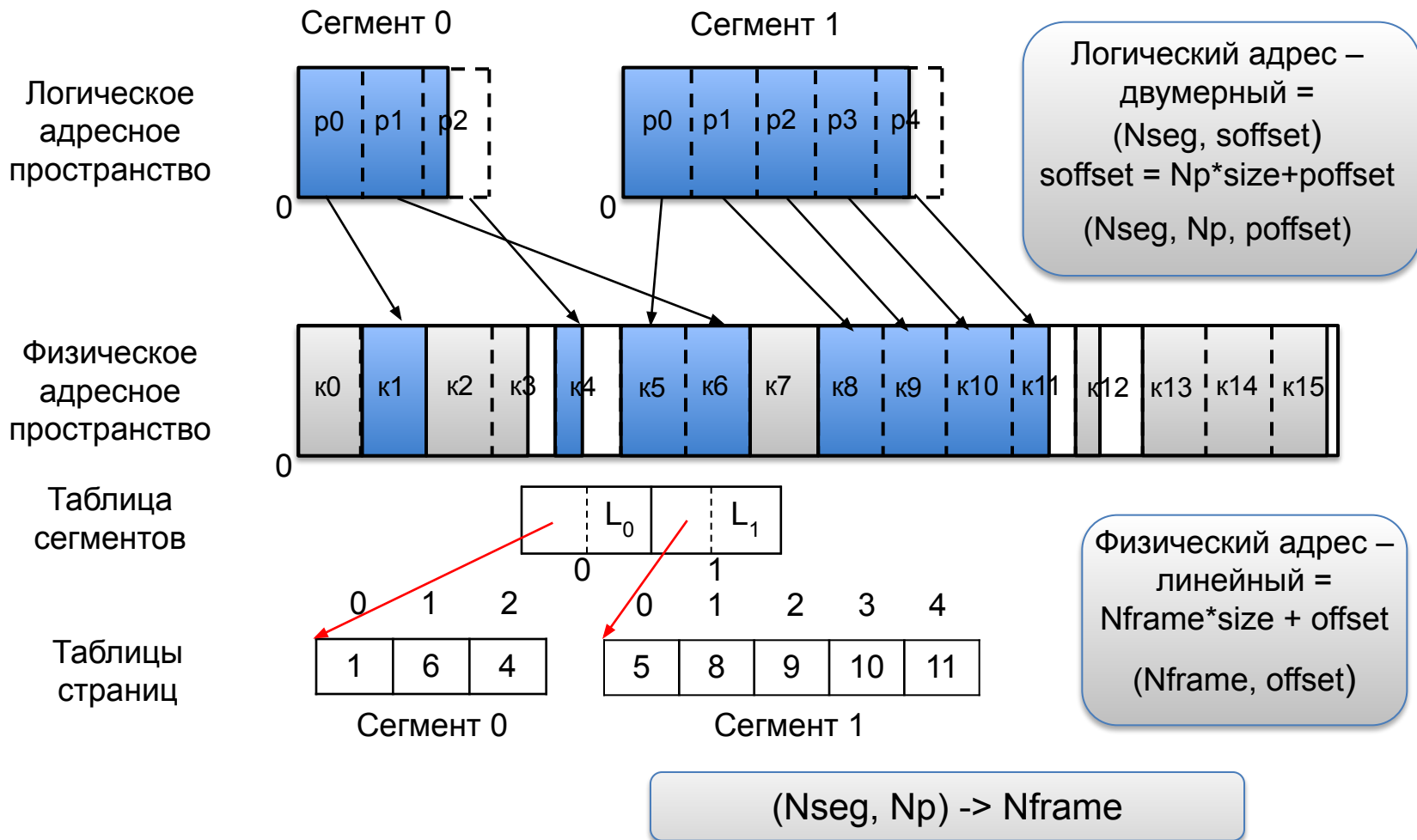
Кусочно-непрерывное отображение

Сегментная организация памяти



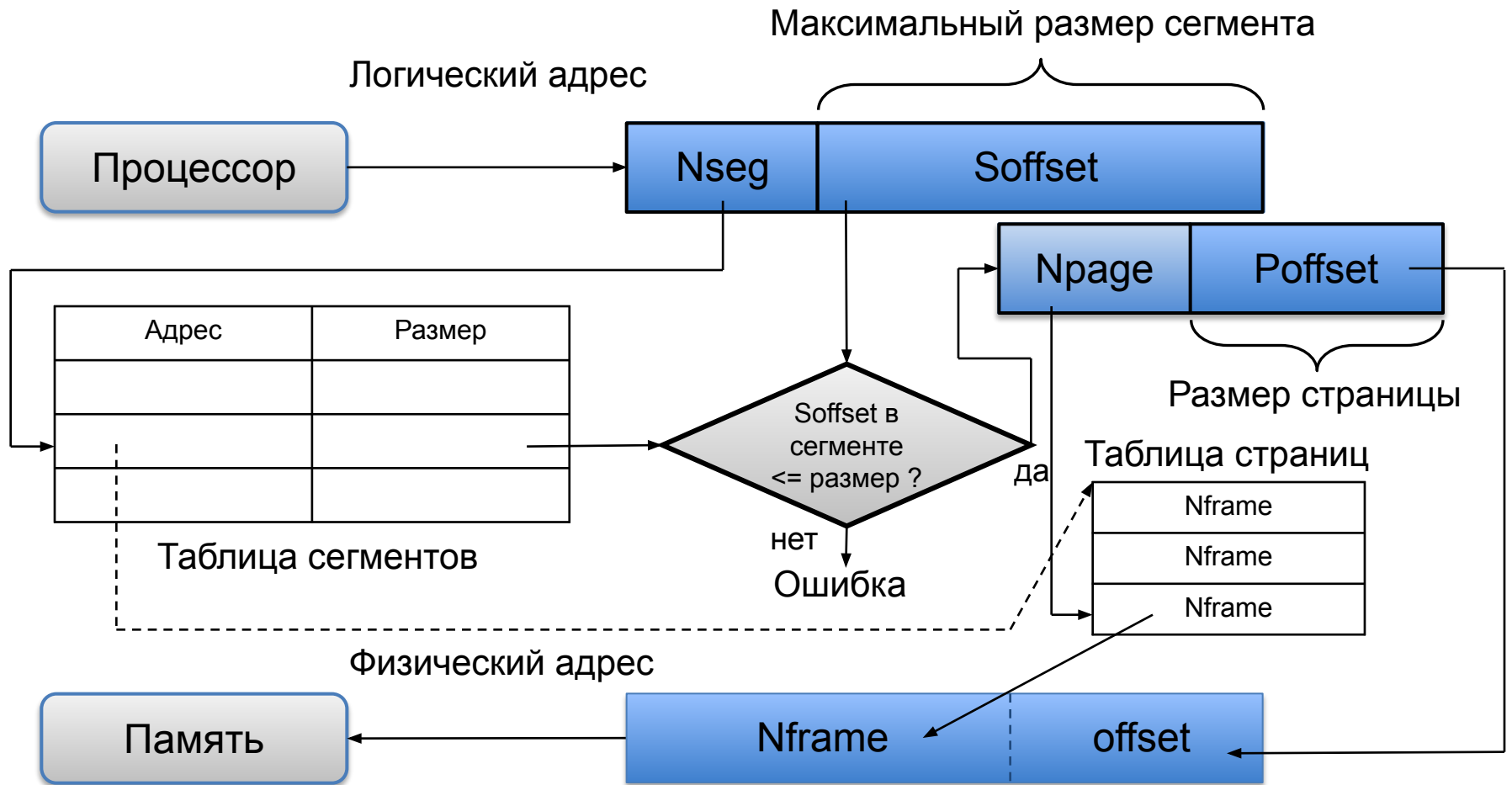
Кусочно-непрерывное отображение

Сегментно-страничная организация



Кусочно-непрерывное отображение

Сегментно-страничная организация



Кусочно-непрерывное отображение

Многоуровневая таблица страниц

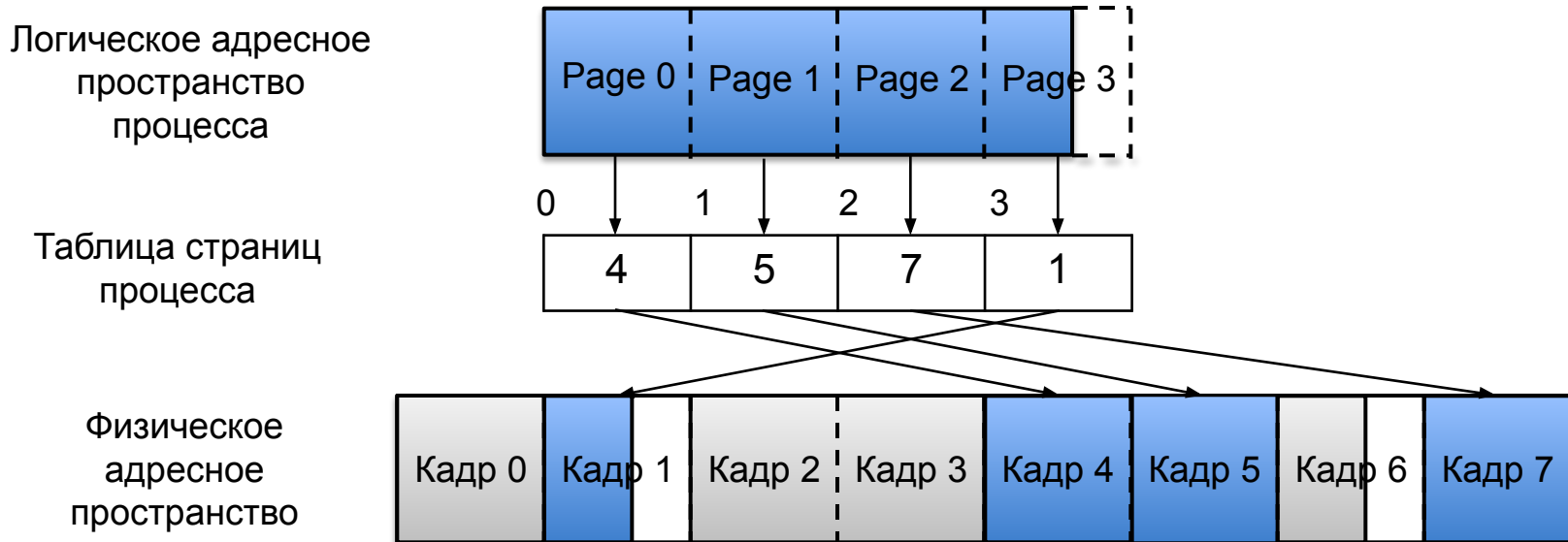
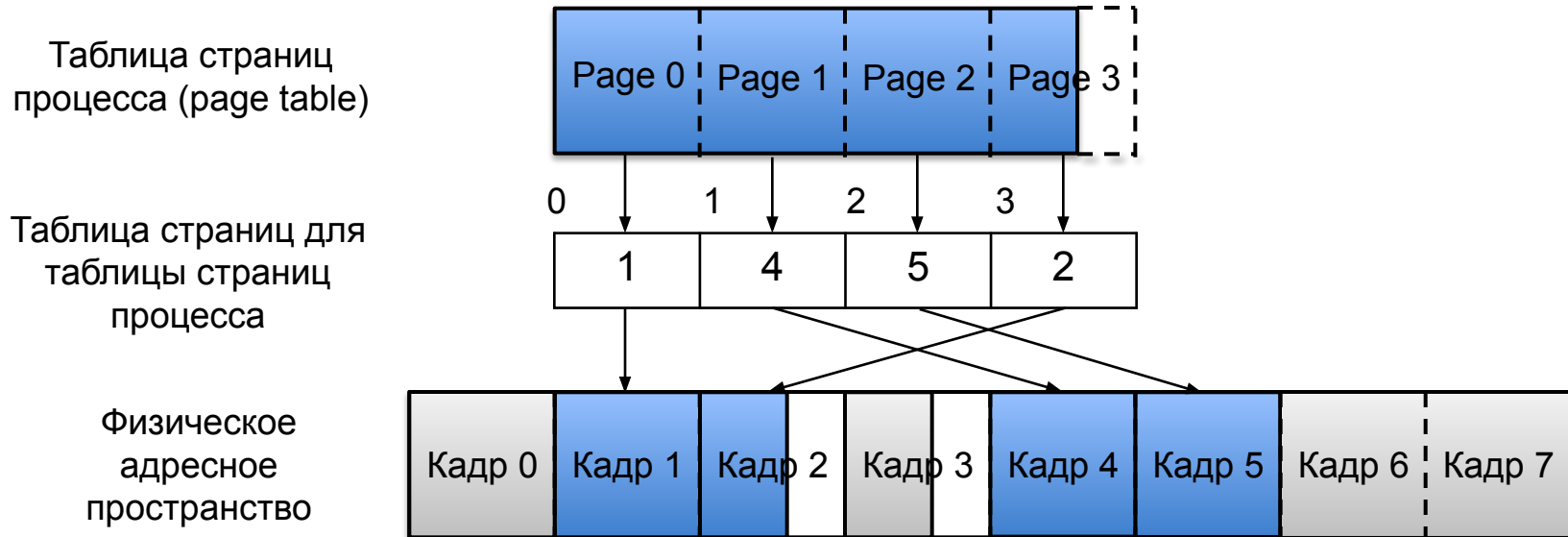


Таблица страниц процесса располагается в физическом адресном пространстве.

При больших размерах таблицы страниц её размещение в последовательных кадрах памяти проблематично.

Кусочно-непрерывное отображение

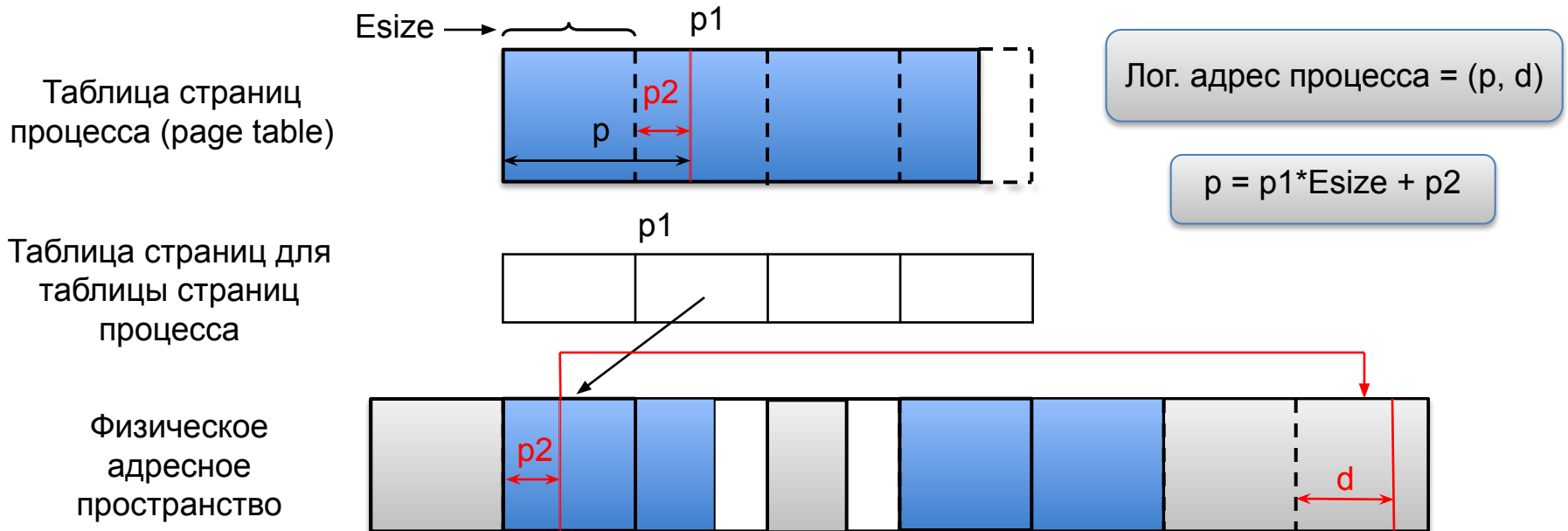
Многоуровневая таблица страниц



Возникает двухуровневая таблица страниц.

Кусочно-непрерывное отображение

Многоуровневая таблица страниц

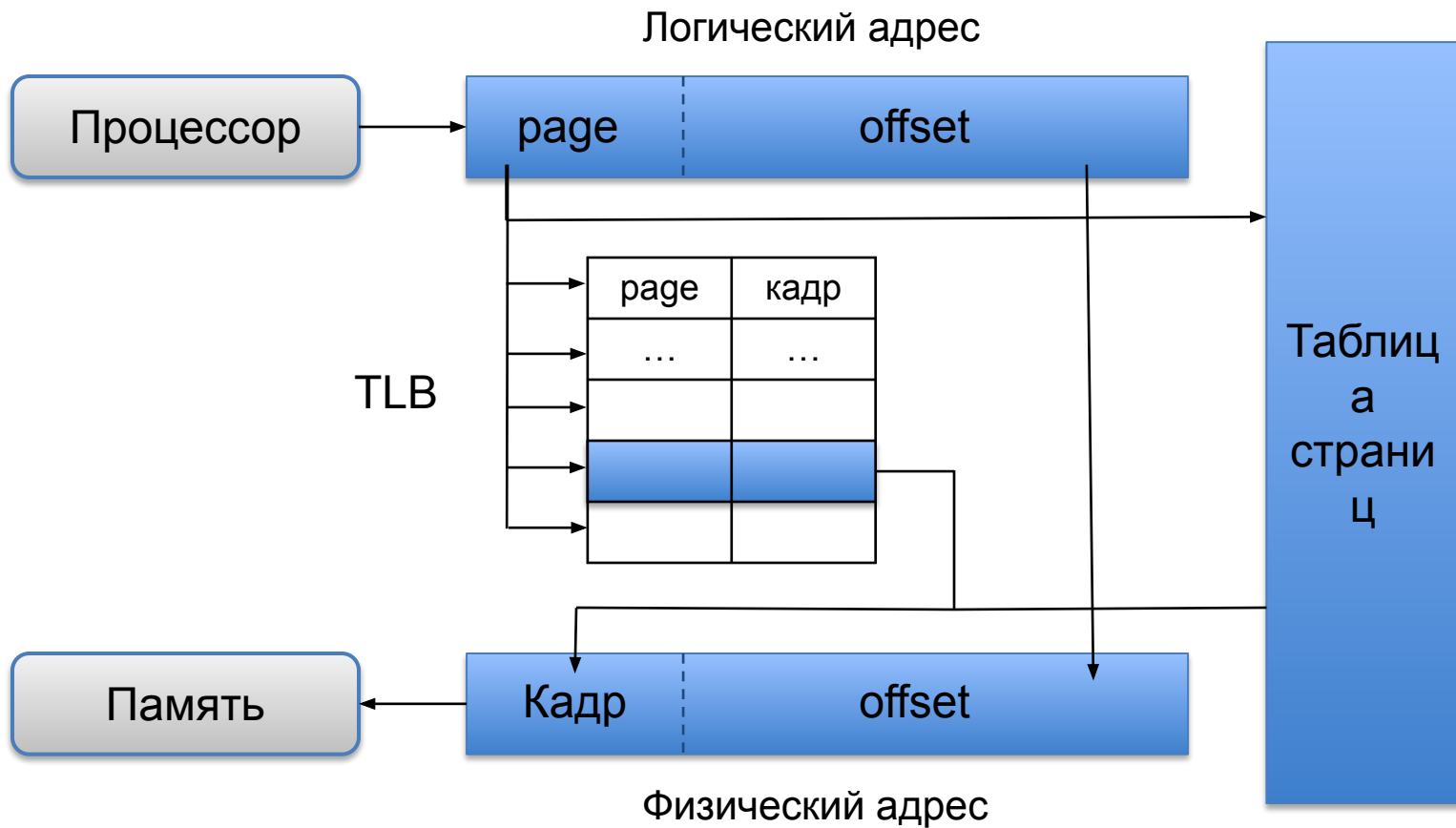


При двухуровневой организации таблицы страниц логический адрес процесса описывается тройкой (p1, p2, d)



Кусочно-непрерывное отображение

Ассоциативная память (TLB)



Кусочно-непрерывное отображение

Ассоциативная память (TLB)

Расчет среднего времени доступа к данному

Обозначения:

t_0 – среднее время доступа к оперативной памяти

t_1 – среднее время доступа к TLB

h – вероятность наличия информации в TLB (hit ratio)

Среднее время доступа к данному при двухуровневой страничной схеме – это:

$$T = t_1 + h*t_0 + (1-h)*3t_0$$

