

# Алгоритмы во внешней памяти

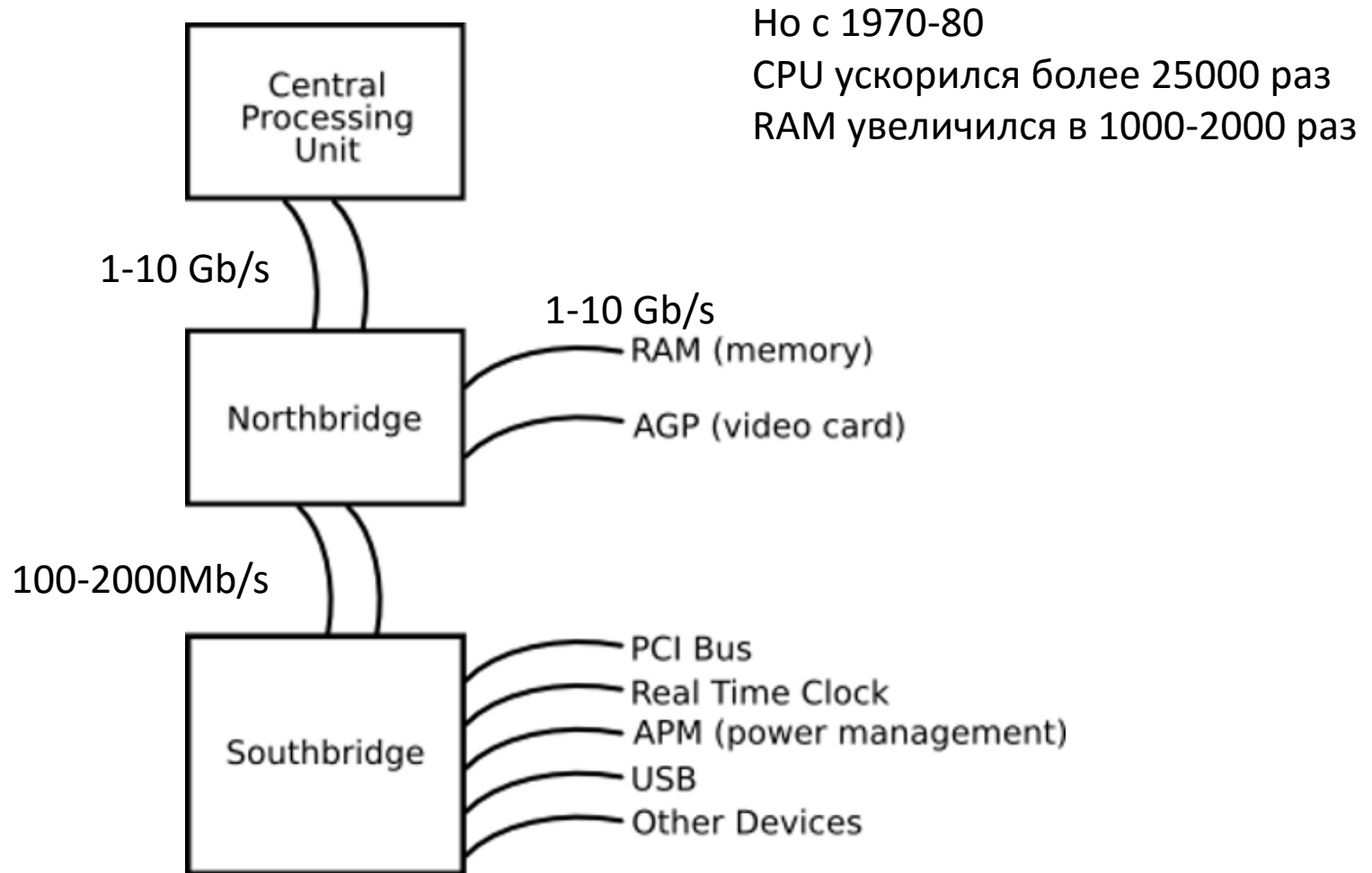
# Другая модель

- Подсчет операций CPU → подсчет операций I/O
- Чтение из RAM ~50-100ns
- Чтение с HDD ~1-10ms
- $N$  – размер входных данных
- $M$  – размер оперативной памяти
- $B$  – размер блока чтения и записи данных
- $T(N, B)$  {по факту  $T(N, M, B)$ } – количество операций чтения/записи при решении задачи

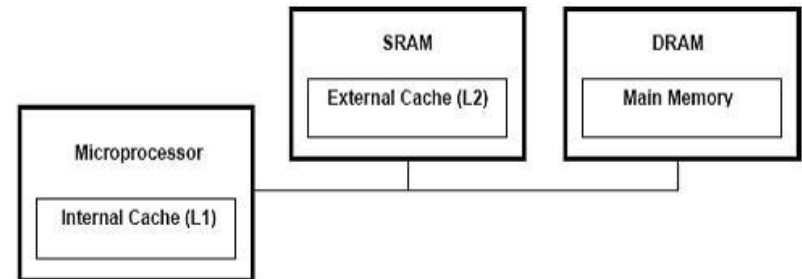
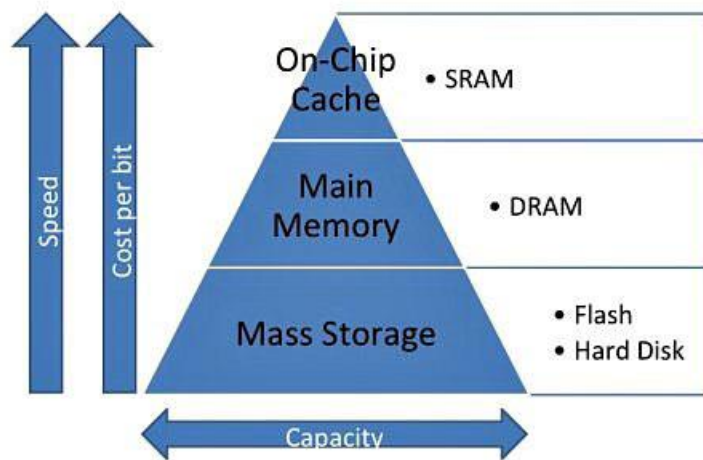
# Пример

- В файле записаны целые числа, нужно найти их сумму
- Завести счетчик в памяти
- Прочитать входные данные и просуммировать
- Вывести результат
- $T(N) = O(N/B) = O(\text{scan})$

# Архитектура



# Оперативная память



# Но кроме нас ...

- На каждом уровне есть буферы
  - В оперативной памяти (у менеджера)
  - На HDD (после seek у себя в буфер он скорее всего скопирует больше, чем вы запросили сразу)
  - У операционной системы (протяжка между файлами и памятью)

# Вопрос?

- Определить скорость доступа к оперативной памяти
- Определить скорость чтения из оперативной памяти
- Определить скорость доступа к HDD/SSD
- Определить скорость чтения с HDD/SSD

# Какие будут задания

- Теоретические задачи
  - Нужно придумать алгоритм (корректный)
  - Подсчитать количество операций I/O
- Практические задачи
  - Реализовать придуманный алгоритм
  - Протестировать его при различных значениях N, M, B
  - Протестировать работу вашего алгоритма в системе автоматического тестирования iRunner
  - Использовать или не использовать существующую библиотеку для работы во внешней памяти (<http://stxxl.sourceforge.net/>)



# Задача 1. Транспонирование

$$A = \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1n} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ a_{m1} & \cdot & \cdot & \cdot & a_{mn} \end{bmatrix} \quad A^H = \begin{bmatrix} \bar{a}_{11} & \cdot & \cdot & \cdot & \bar{a}_{m1} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \bar{a}_{1n} & \cdot & \cdot & \cdot & \bar{a}_{mn} \end{bmatrix}$$

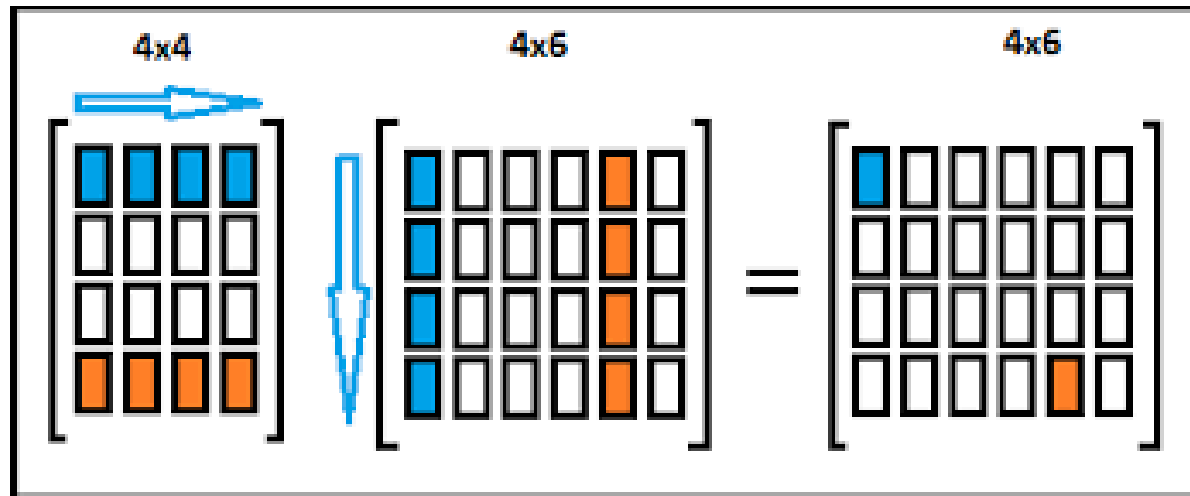
1)  $O(\text{scan})$

1)  $B^2 < M$

2)  $n, m > B$

2)  $O(N/\min(B, M^{0.5}))$

# Задача 2. Перемножение матриц



Пусть матрицы квадратные  $K \times K$

- 1)  $O(K^3/B)$
- 2)  $O(K^3/B^2)$ ,  $B^2 < M$
- 3)  $O(K^3/\min(B^2, B \times M^{0.5}))$

# Множества

- $P = \{p_1, p_2, \dots, p_n\}$
- $Q = \{q_1, q_2, \dots, q_m\}$
- $P \cap Q = ?$
- $P \cup Q = ?$
- $P/Q = ?$
- Все операции  $O(\text{sort})$ .

# Порядковая статистика

- $A = \{a_1, a_2, \dots, a_n\}$
- Найти медиану  $A$ :
  - $b_1 \leq b_2 \leq \dots \leq b_{\{n/2\}} \leq \dots \leq b_n$   
(перестановка элементов  $A$ )
  - Найти  $b_{\{n/2\}}$ .