

Московский Государственный Университет им. М.В. Ломоносова
Факультет Вычислительной Математики и Кибернетики
_____Кафедра Суперкомпьютеров и Квантовой Информатики_____



**Спецкурс: системы и средства параллельного
программирования**

Отчёт № 2

**Анализ влияния кэша на операцию блочного
матричного умножения**

Работу выполнила
Домрачева Д. А.

Постановка задачи и формат данных

Задача:

Реализовать последовательный алгоритм матричного умножения и оценить влияние кэша на время выполнения программы.

Формат командной строки:

<имя файла матрицы A> <имя файла матрицы B> <имя файла матрицы C> <режим, порядок индексов>.

Параметры командной строки, определяющие режим выполнения:

- 0 – ijk , 1 – ikj ;
- 32 – блок размера 32×32 , или 0 – блок оптимального размера, вычисляемый по формуле: $3 \times b^2 = m \times L$, где b – размер блока в элементах, а $m \times L$ – размер кэша;
- 1, 2, 3, 4, 5, 6, 7 – режимы для вычисления определенных параметров системы, могут быть заданы в произвольном порядке, где, соответственно:
 - 1 – RAPI_L1_TCM – промахи кэша L1;
 - 2 – RAPI_L2_TCM – промахи кэша L2;
 - 3 – RAPI_L3_TCM – промахи кэша L3;
 - 4 – RAPI_TOT_INS – общее число выполненных инструкций;
 - 5 – RAPI_TLB_IM – промахи инструкций TLB;
 - 6 – RAPI_TOT_CYC – общее число циклов;
 - 7 – RAPI_TLB_DM – промахи данных TLB.

Примечание: среди доступных счетчиков на устройстве, где выполнялись запуски программы, нет счетчиков для определения числа FLOP.

Формат файла с матрицей:

Матрица представляются в виде бинарного файла следующего формата:

Тип	Значение	Описание
Число типа $size_t$	N – натуральное число	Число строк матрицы
Число типа $size_t$	M – натуральное число	Число столбцов матрицы
Массив чисел типа T	$N \times M$ элементов	Массив элементов матрицы

Элементы матрицы хранятся построчно.

Описание алгоритма

Математическая постановка:

Алгоритм матричного умножения ($A \times B = C$) можно представить в следующем виде:

$$c_{ij} = \sum_{k=1}^n \sum_{k_1=k}^{k+b} (a_{ik_1} \cdot b_{k_1j}) \quad \text{для каждого элемента матрицы } C.$$

Оценка влияния кэша на время выполнения программы осуществляется за счёт перестановки индексов суммирования и за счет изменения размеров блока.

Анализ выполнения:

Для оценки времени выполнения программы использовалась функция *clock()*. Для оценки таких параметров, как промахи кэша, промахи TLB и количество процессорных тактов, использовались счетчики библиотеки *rapi.h*.

Верификация:

Для проверки корректности работы программы использовались тестовые данные. Для нахождения среднего времени выполнения умножения матриц программа запускалась по 5 раз.

Основные функции:

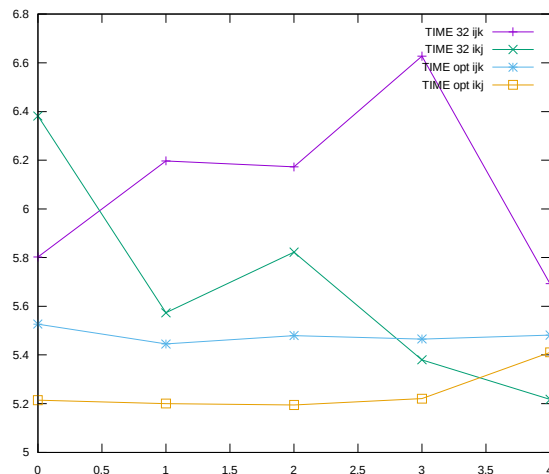
- Выделение памяти для матриц – *init_matrix*. Чтение матриц из бинарных файлов – *read_matrix*. Функция для вывода матриц в файл – *write_matrix*.
- Основная функция, определяющая время выполнения умножения – *get_time*. Вызывает дополнительную функцию, реализующую умножение – *multiply_ijk* или *multiply_kj*, в зависимости от заданного режима перемножения.
- Функции, реализовывающие умножение матриц в зависимости от порядка индексов и размера блоков – *multiply_ijk* и *multiply_kj*. Возвращают количество тактов, за время которых выполнялись операции.

Результаты выполнения

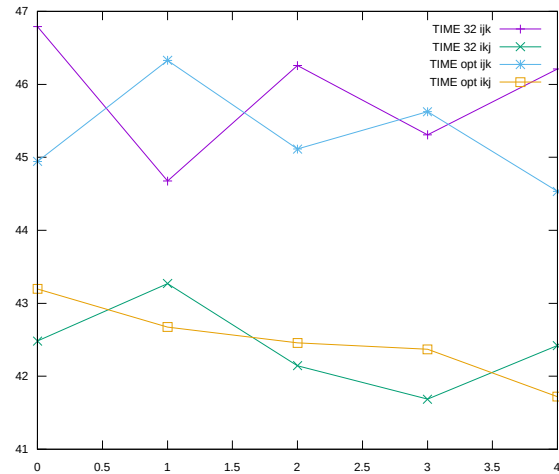
Результаты:

Проводилось перемножение матриц размерами 1000×1000 и 1000×1000 , 2000×2000 и 2000×2000 , 3000×3000 и 3000×3000 , 4000×4000 и 4000×4000 , 5000×5000 и 5000×5000 . Зависимость среднего времени выполнения от порядка индексов суммирования и размера блоков представлена на графике (время в секундах).

1000×1000 :



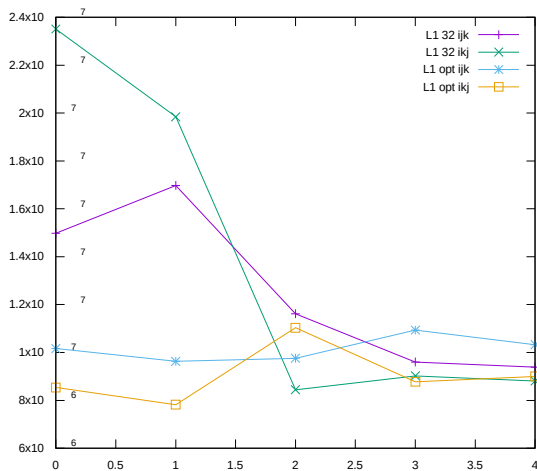
2000×2000 :



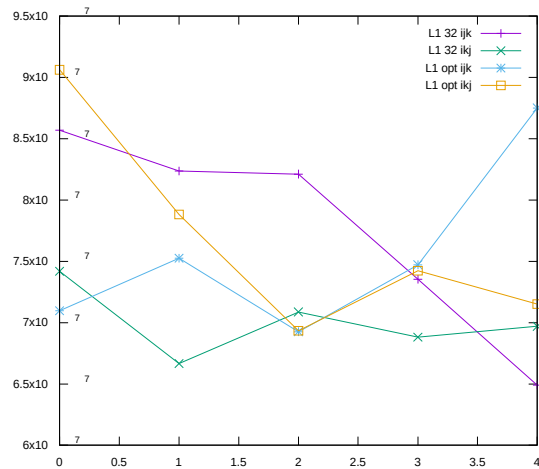
3000×3000 :

Зависимость промахов кэша L1.

1000×1000 :

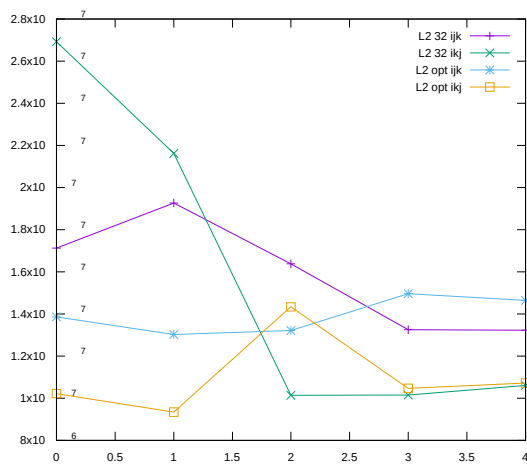


2000×2000 :

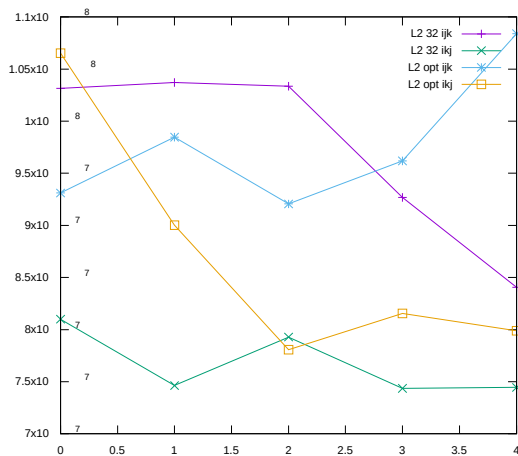


Зависимость промахов кэша L2.

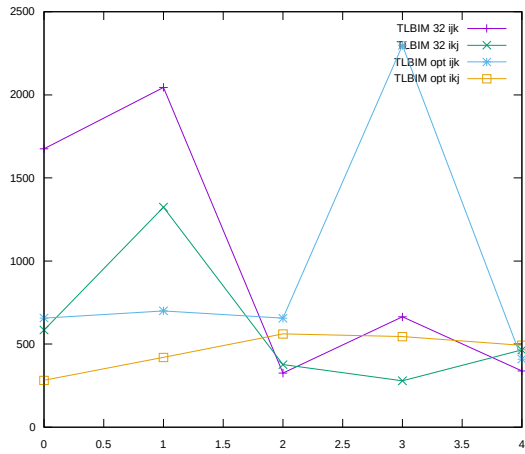
1000×1000 :



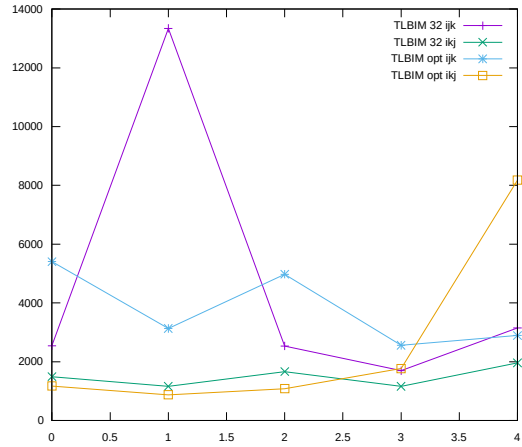
2000×2000 :



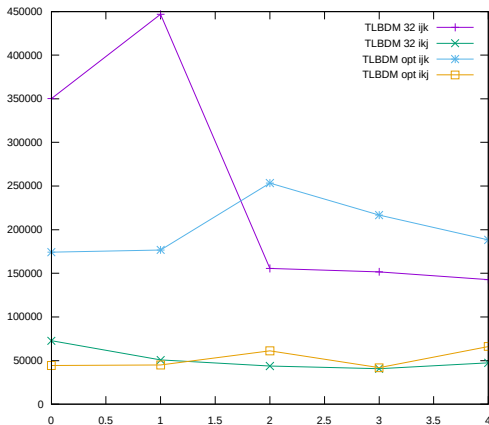
Зависимость промахов инструкций TLB.
1000×1000 :



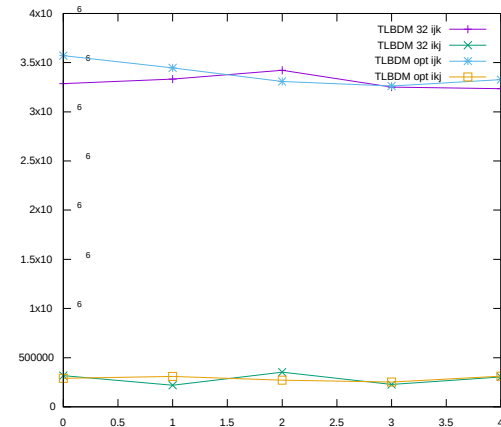
2000×2000 :



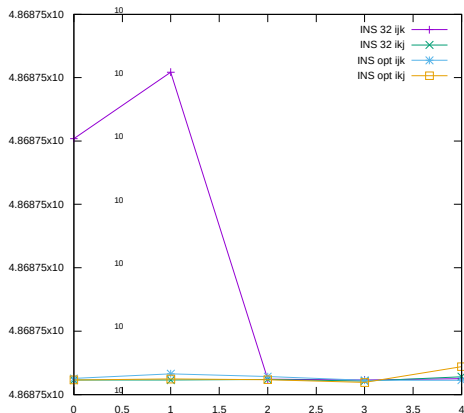
Зависимость промахов данных TLB.
1000×1000 :



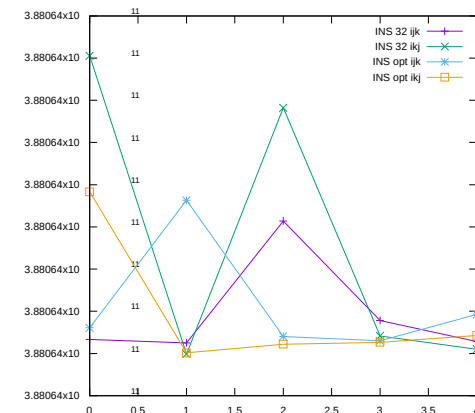
2000×2000 :



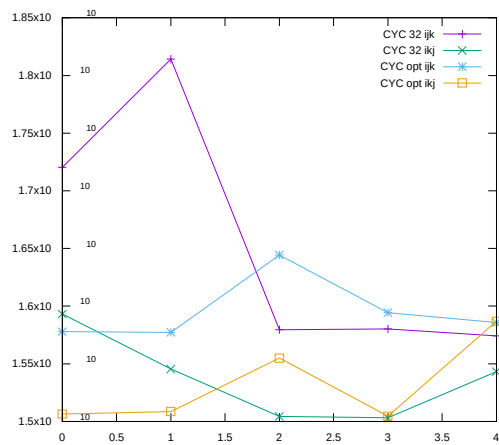
Зависимость количества инструкций.
1000×1000 :



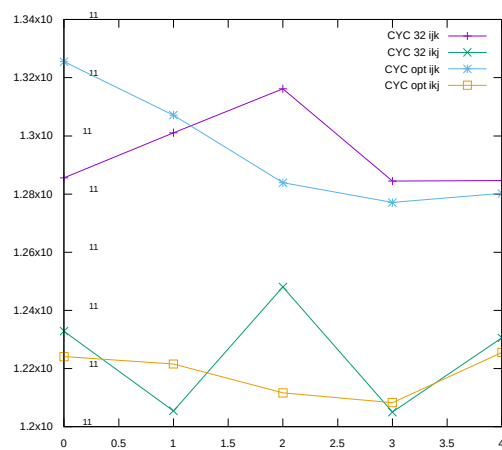
2000×2000 :



Зависимость количества циклов.
1000×1000 :



2000×2000 :



Основные выводы