



**УНИВЕРСИТЕТ
ЛОБАЧЕВСКОГО**



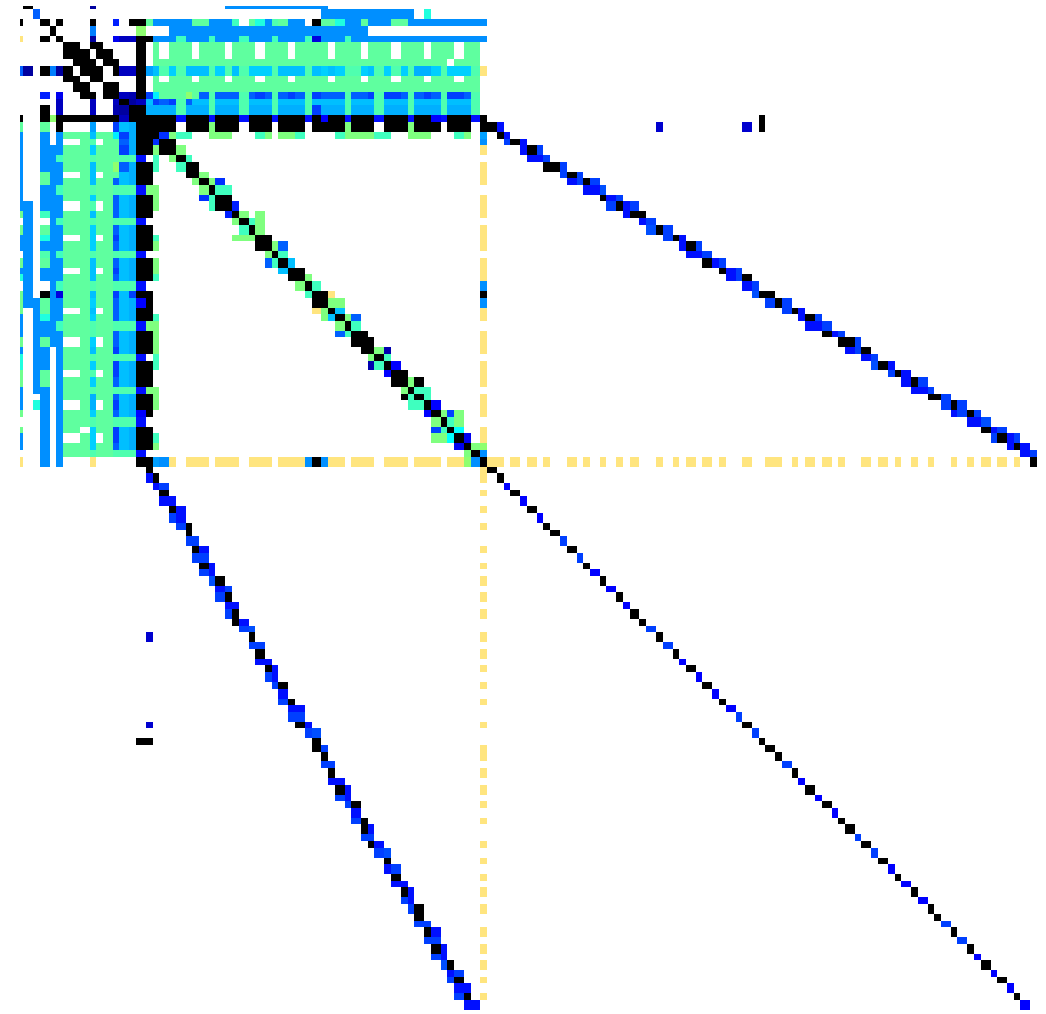
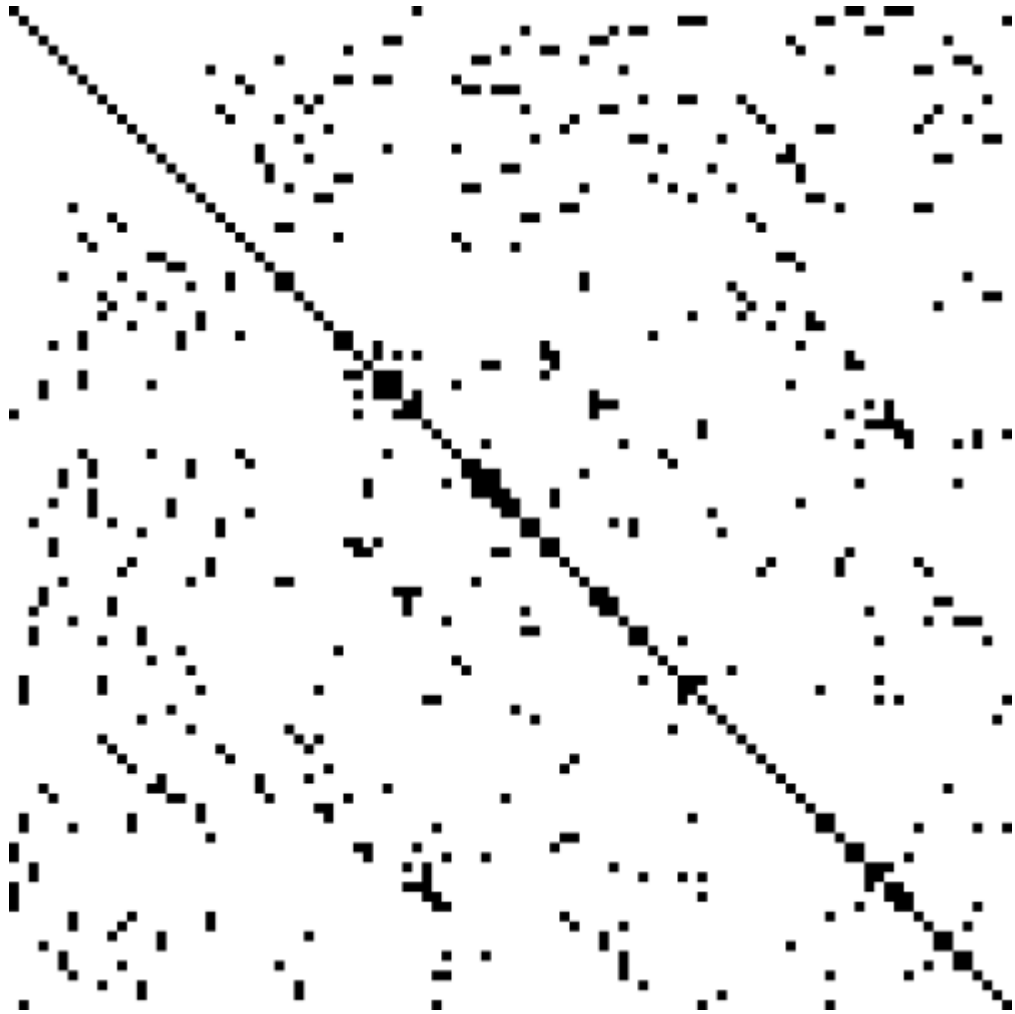
**НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. Н.И. ЛОБАЧЕВСКОГО
ИНСТИТУТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МАТЕМАТИКИ И МЕХАНИКИ
КАФЕДРА ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛЕНИЙ И
СИСТЕМНОГО ПРОГРАММИРОВАНИЯ**

Новый формат разреженных матриц

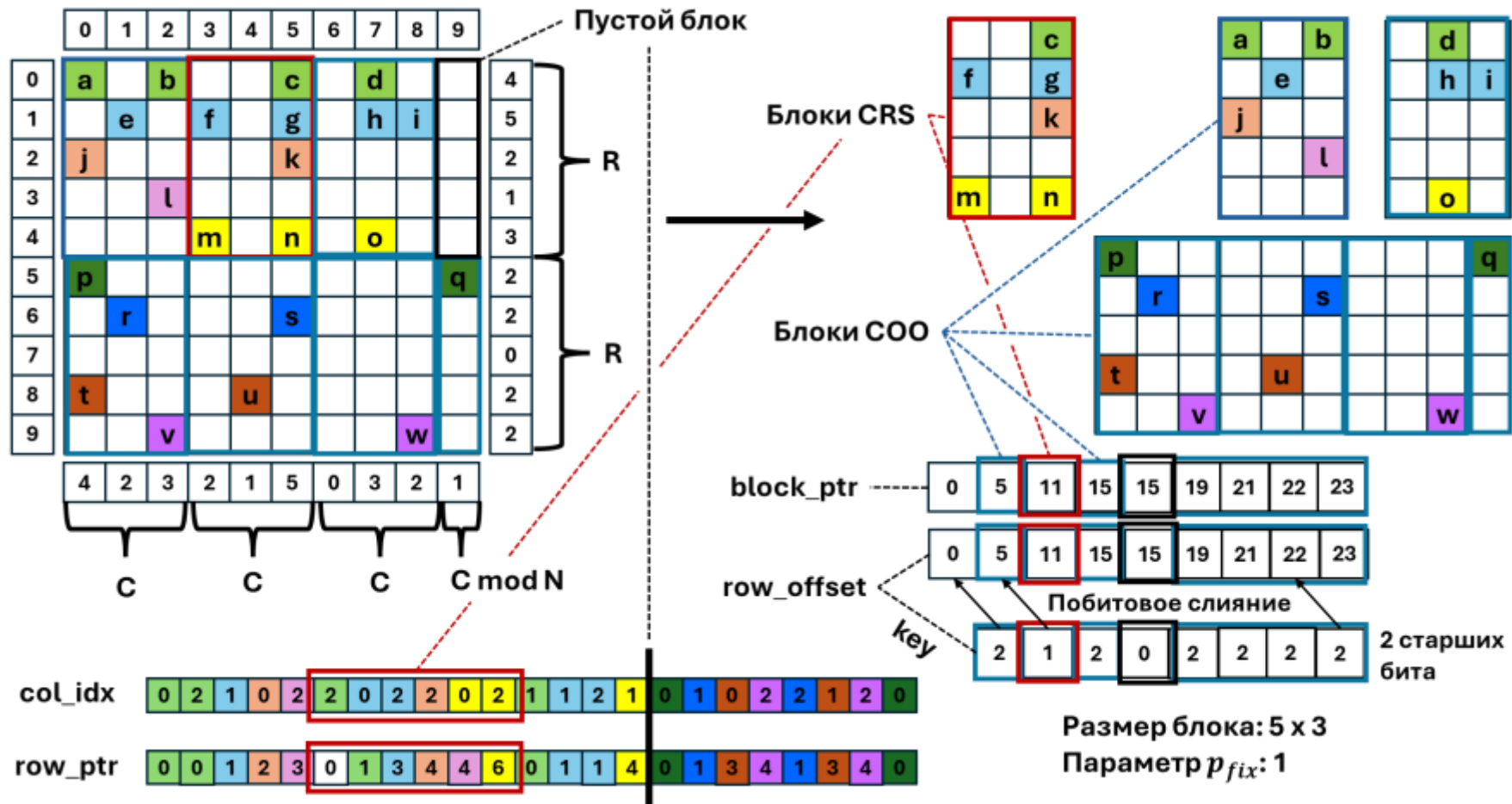
М.А. Загрядсков, В.Д. Волокитин

«Суперкомпьютерные дни в России»
Москва, 28-29 сентября 2026 г.

Разреженные матрицы, SpMV ($y := \alpha Ax + \beta y$)



- ❑ **Основная идея:** улучшение локальности обхода **вектора x**
- ❑ Исходная матрица разбивается на блоки в различных подформатах



Вычислительные эксперименты

❑ Тестовое окружение:

- плата Banana PI BPI-F3, процессор SpacemiT Keystone K1, in-order, 8 core SpacemiT x60, 1.6 GHz, 1MB разделённого кэша L2, 256-bit RVV 1.0, 16 GB, OC Bianbu 2.2.1
- плата Banana PI BPI-SM10, процессор SpacemiT Keystone K3, OoO, 8 core SpacemiT x100, 2.4GHz, 2MB разделённого кэша L2, 256-bit RVV 1.0, 8 GB, OC Bianbu 4.0.1
- Кросс-компилятор GCC RISC-V 14.2.0

❑ Тестовая выборка:

- 121 матрица из коллекции SuiteSparse размером от $0.5 \cdot 10^3 \times 400 \cdot 10^3$ до $8.39 \cdot 10^6 \times 8.39 \cdot 10^6$ и заполненностью от $3.1 \cdot 10^{-7}$ до $1.4 \cdot 10^{-2}$. Плотный вектор заполнялся случайными числами из интервала $[0, 1)$, значения $\alpha = 1.0$, $\beta = 0.0$ фиксированы

❑ Тестовые параметры:

- наилучшие значения параметров R, C, p_{fix} подобраны экспериментально для каждой из тестовых платформ и использовались одни и те же для всех матриц в эксперименте

❑ Мера производительности:

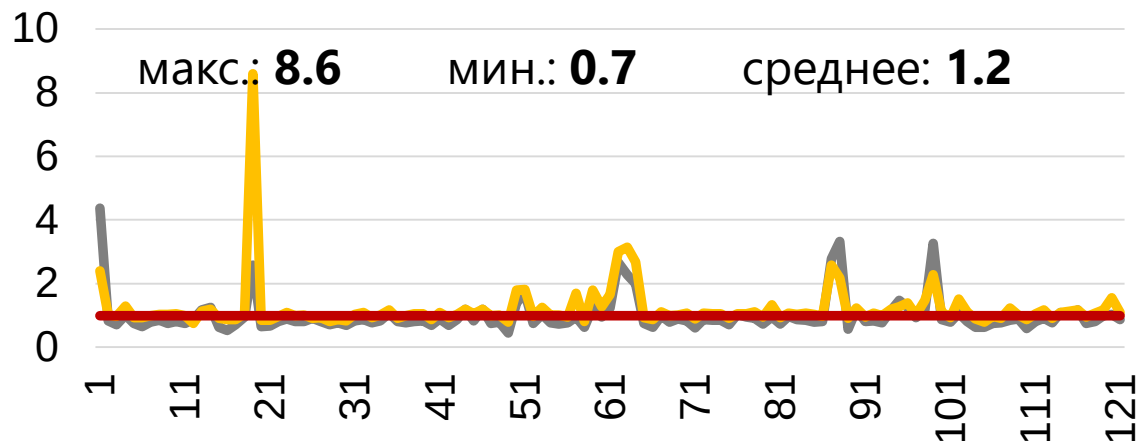
- относительное ускорение, отношение времени работы **SpMV** в формате **CSR** к времени работы в формате **HCSR**



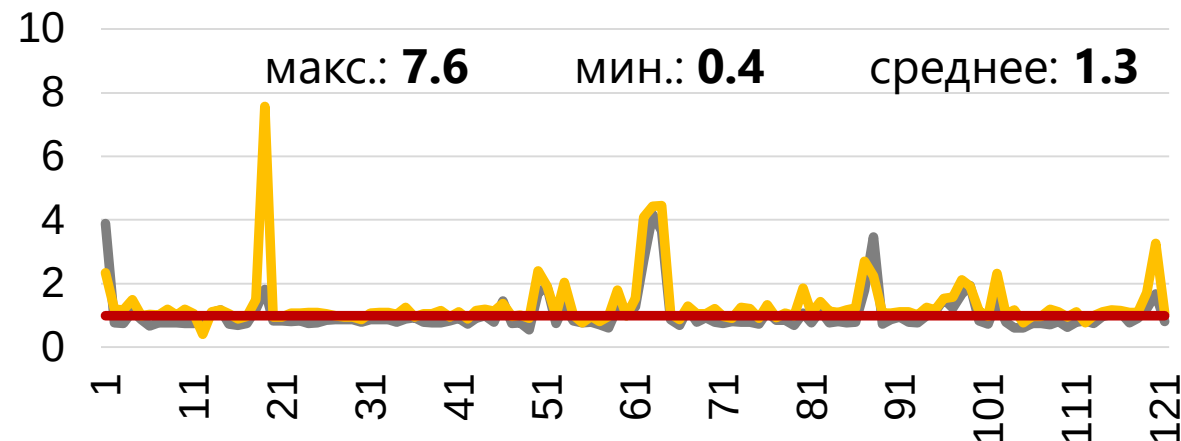
Вычислительные эксперименты

Относительное ускорение

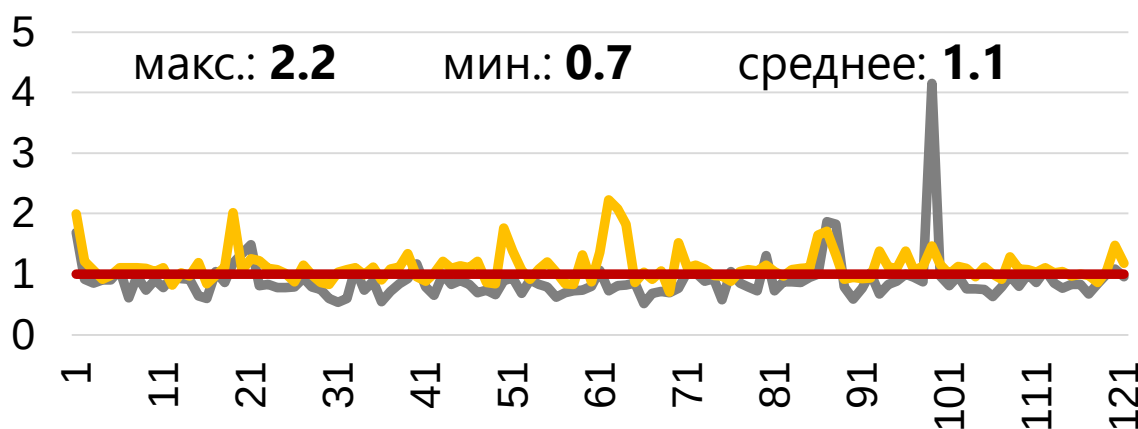
SpacemiT K1, double



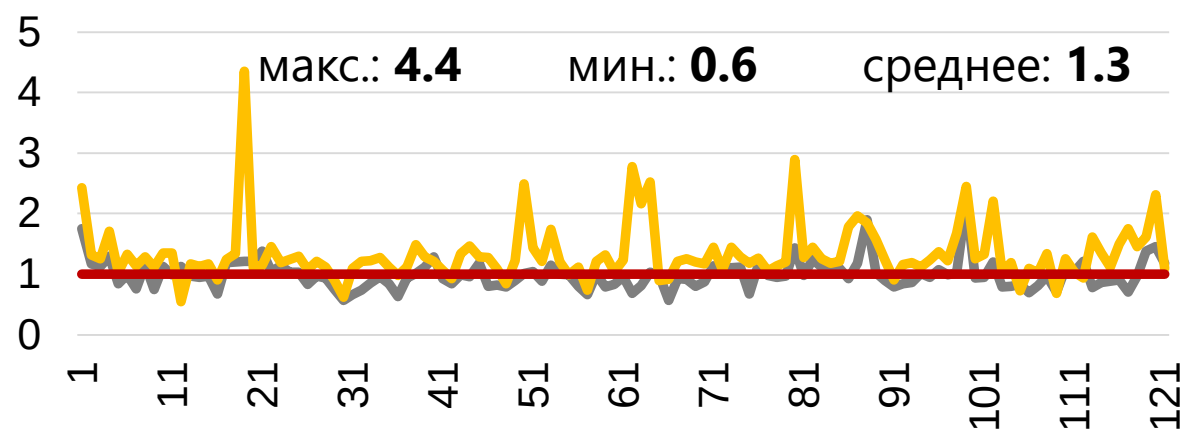
SpacemiT K1, float



SpacemiT K3, double



SpacemiT K3, float

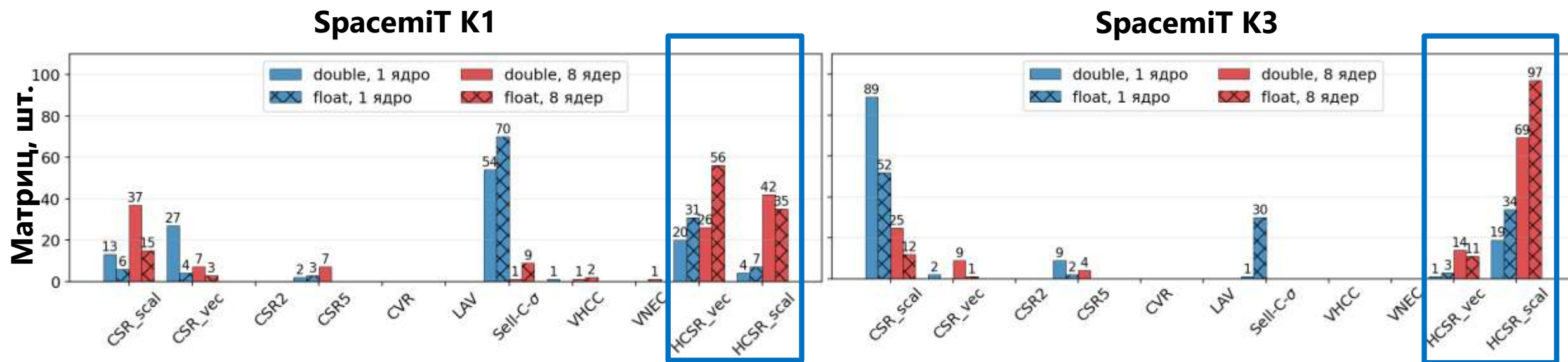


ID матрицы



Вычислительные эксперименты

Число матриц, на которых разреженный формат показывает лучший результат (Тор-1)



Выводы

- ❑ **CSR** и **HCSR** показывают наилучшее время на наибольшем количестве ядер (8)
- ❑ Ускорение от CSR до **8.6 раз** в double и до **7.6 раз** в float на **8 ядрах**
- ❑ Прирост на **1 ядре** при векторизации на **SpacemiT K1**
- ❑ Чувствительность **HCSR** к смене подсистемы памяти и конвейера в **SpacemiT K3**: максимальное ускорение снизилось, среднее ускорение почти не изменилось.
- ❑ Преимущество скалярного **HCSR** на **SpacemiT K3**: лучший результат на **наибольшем** числе матриц в сравнении с другими форматами на **8 ядрах**.
- ❑ Преимущество скалярного и векторизованного **HCSR** на **SpacemiT K1** в сравнении с другими форматами на **8 ядрах**.
- ❑ Реализации всех форматов доступны на **GitHub**:
<https://github.com/UNN-ITMM-Software/RVVLASparse>

