

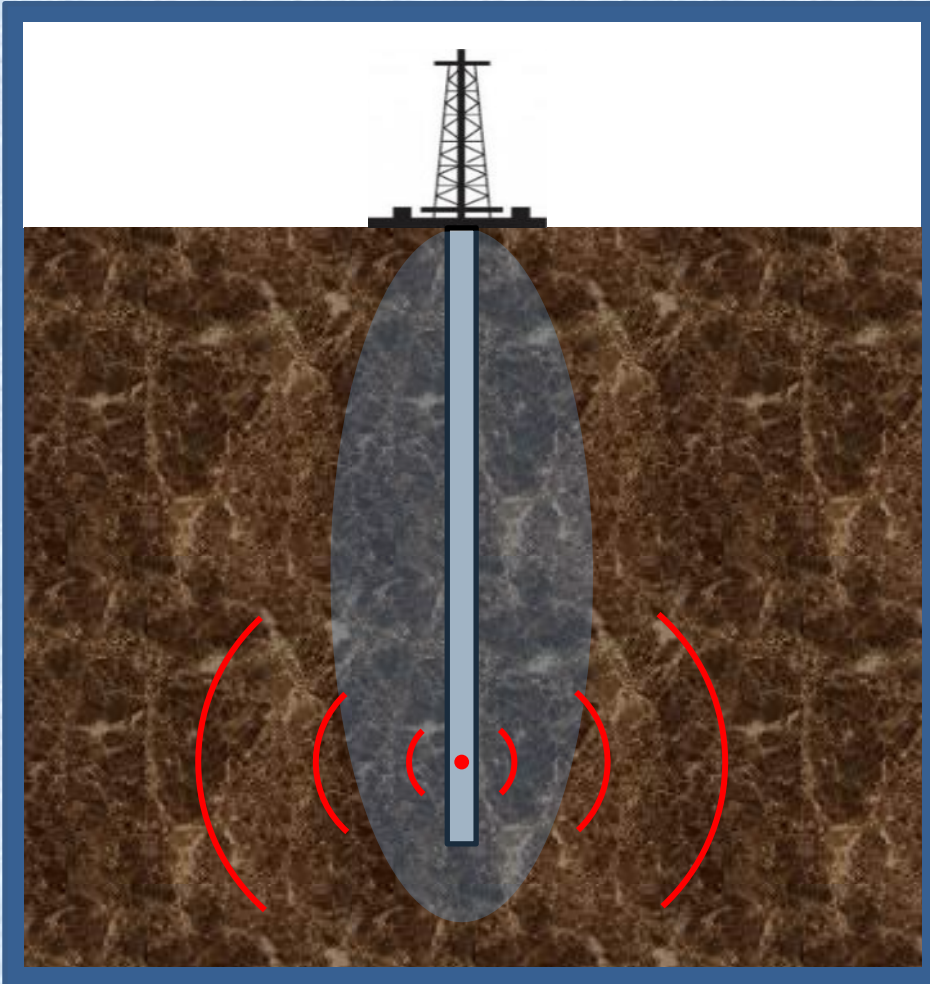
**Сравнение производительности кластеров
НКС-1П (ССКЦ СО РАН) и МВС-10П (МСЦ РАН)
на задаче моделирования
нелинейной многофазной фильтрации
в деформируемых пористых средах**

**Киреев С.Е.
ИВМиМГ СО РАН**

План доклада

- **Описание задачи**
 - Введение
 - Математическая модель
 - Численный метод
 - Алгоритм
 - Параллельная реализация
- **Описание кластеров**
 - Архитектура
 - Вопросы использования
- **Результаты сравнительного тестирования**

Описание задачи



Объект исследования

- математическое моделирование задач многофазной фильтрации в деформируемых пористых средах на вычислительных системах сверхвысокой производительности

Актуальность

- задачи разработки нефтяных месторождений

Предыстория

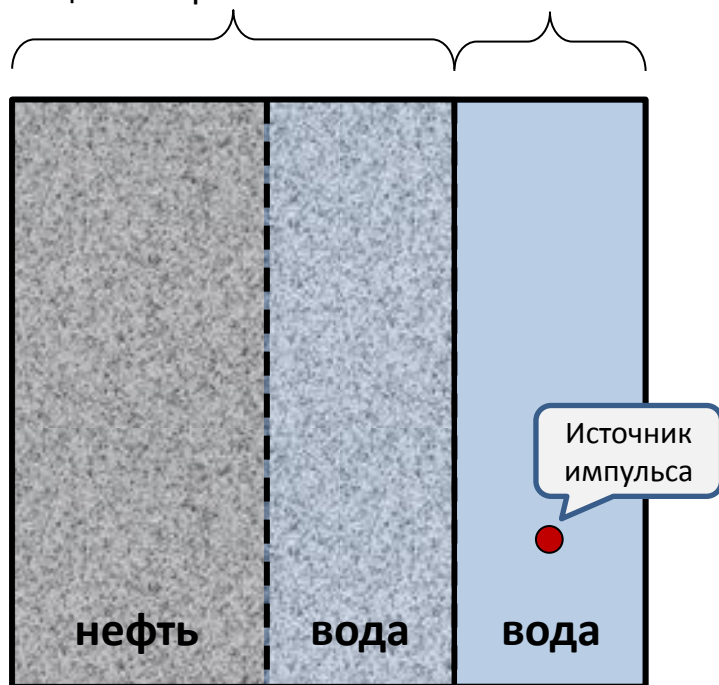
- Была разработана модель, которая позволяет описывать нелинейную неизотермическую фильтрацию водонефтяной смеси сквозь упруго-деформируемую пористую среду
 - E.I. Romenskiy, Y.V. Perepechko, G.V. Reshetova «Modeling of Multiphase Flow in Elastic Porous Media Based on Thermodynamically Compatible Systems Theory» // Poster presentation at ECMOR XIV - 14th European conference on the mathematics of oil recovery
- На основе модели был создан параллельный программный комплекс *Porodynamics*, предназначенный для использования на кластере с графическими ускорителями Nvidia
 - Государственный контракт № 07.514.11.4156 «Поисковые проблемно-ориентированные исследования в области математического моделирования задач многофазной фильтрации в деформируемых пористых средах на вычислительных системах сверхвысокой производительности».
- **На основе модели был разработан параллельный программный комплекс для кластеров с ускорителями Intel Xeon Phi**
 - С.Е. Киреев Оптимизация для кластера с ускорителями Xeon Phi задачи фильтрации водонефтяной смеси через эластичную пористую среду // Вычислительные методы и программирование, т.16, выпуск 2, 2015, с. 187-195.

Пример расчёта

Распространение волны в пористой формации

Насыщенная пористая
формация с пористостью 0.2

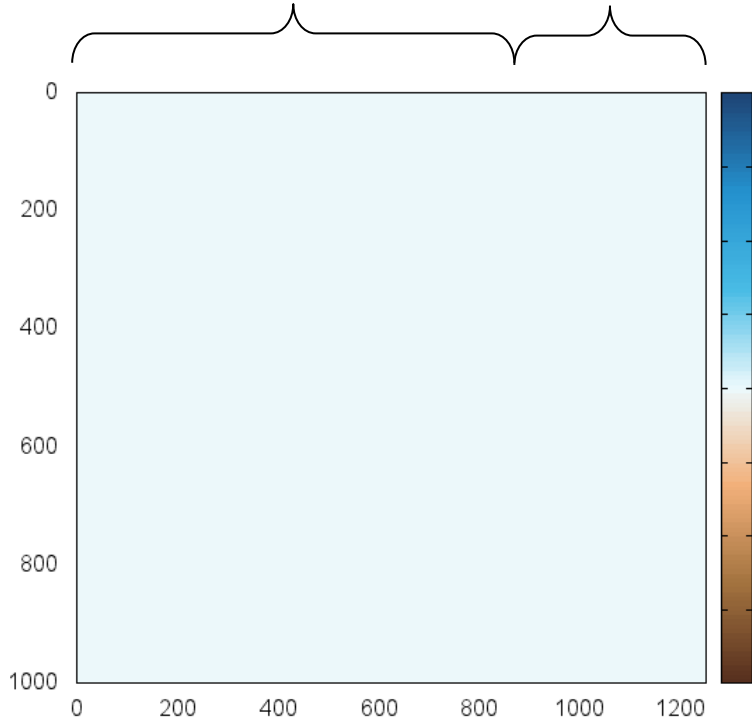
Скважина



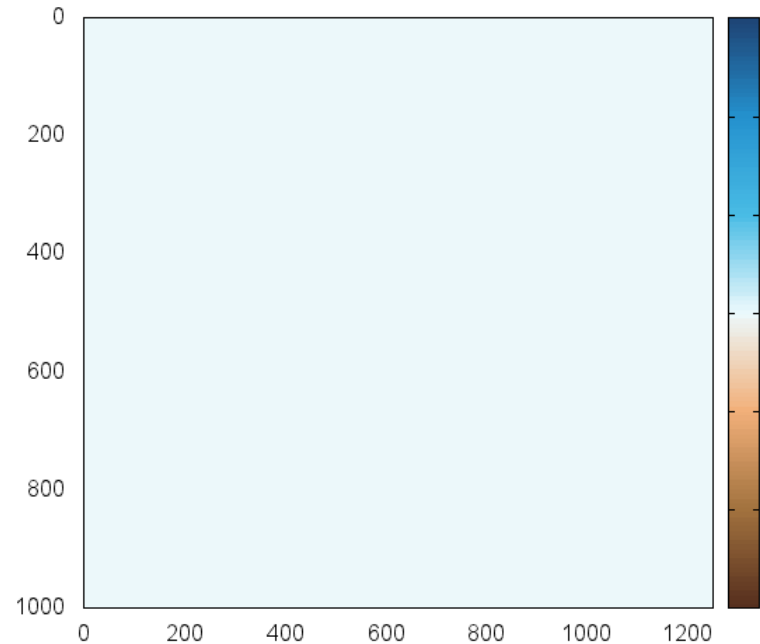
Пример расчёта

Распространение волны в пористой формации

Насыщенная пористая
формация с пористостью 0.2 Скважина



давление

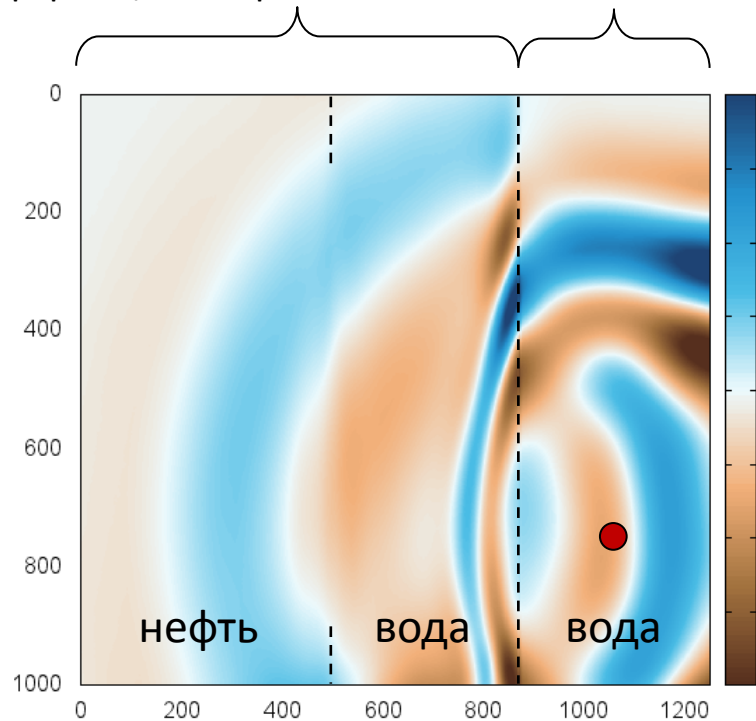


компонента X тензора напряжений
формации

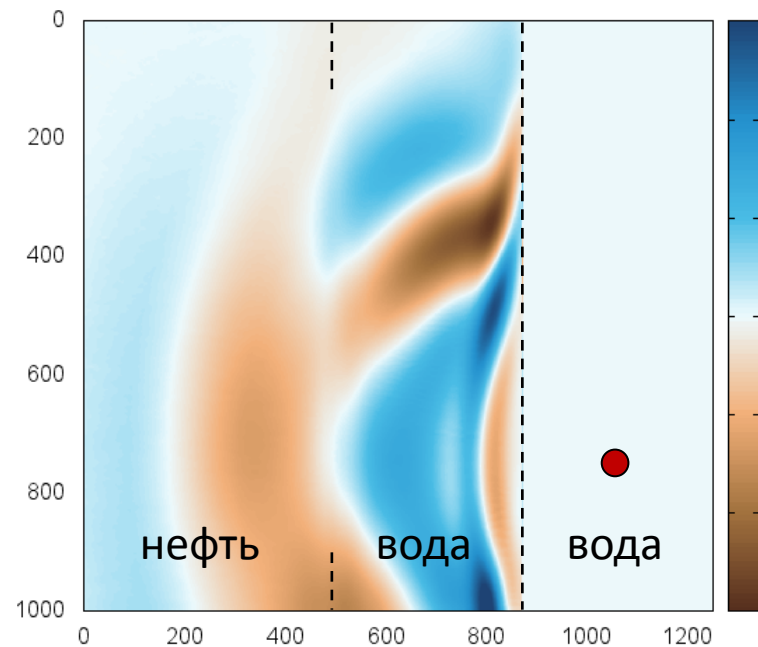
Пример расчёта

Распространение волны в пористой формации

Насыщенная пористая
формация с пористостью 0.2 Скважина



давление



компонента X тензора напряжений
формации

Пример расчёта 2

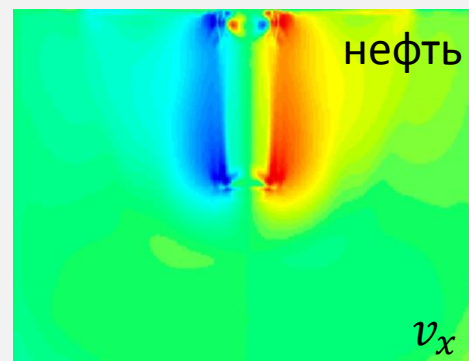
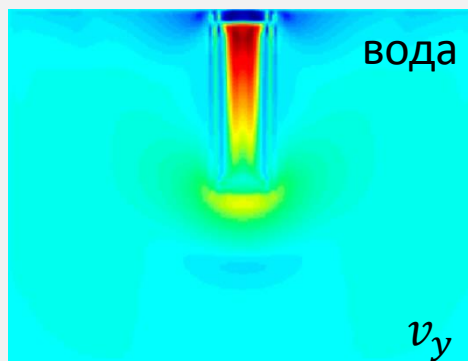
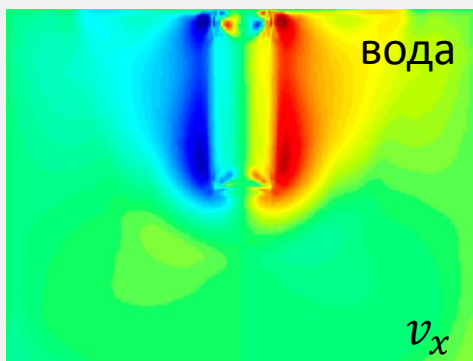
Фильтрация водно-нефтяной смеси



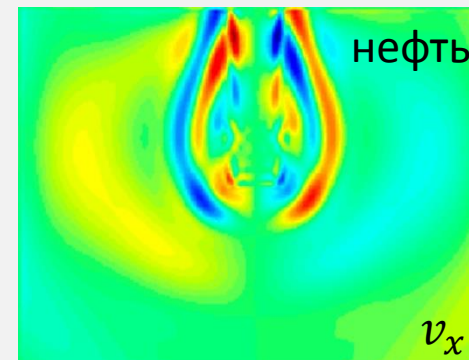
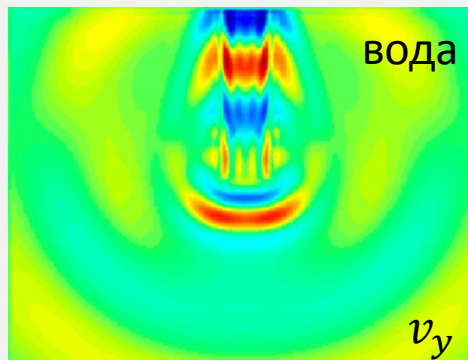
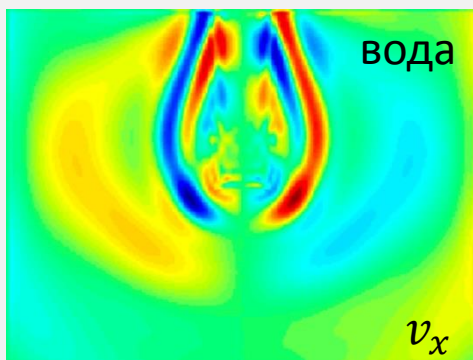
Пример расчёта 2

Фильтрация водно-нефтяной смеси

- Просачивание водно-нефтяной смеси в пространство за пределами шахты, когда смесь закачивается в шахту.



- Поля скоростей воды (горизонтальная (слева) и вертикальная (справа) компоненты) и нефти (горизонтальная компонента). Источник давления отсутствует.



- Поля скоростей воды (горизонтальная (слева) и вертикальная (справа) компоненты) и нефти (горизонтальная компонента). Источник давления, заданный импульсом Рикера, расположен в центре нижней части шахты.

Модель среды

Термодинамически согласованная модель фильтрации многофазной смеси в деформируемой пористой среде

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_k}{\partial x_k} = 0 \qquad \frac{\partial \alpha_n \rho_n}{\partial t} + \partial_k (\alpha_n \rho_n u_k^n) = 0$$

$$\frac{\partial \rho \alpha_n}{\partial t} + \partial_k (\rho \alpha_n u_k) = - \sum_{m=2}^N \lambda_{nm} (p_1 - p_m)$$

$$\frac{\partial w_k^n}{\partial t} + \partial_k \left(\frac{u_i^1 u_i^1}{2} - \frac{u_i^n u_i^n}{2} + e^1 + \frac{p_1}{\rho_1} - e^n - \frac{p_n}{\rho_n} \right) = e_{kij} u_i \omega_j^n - \sum_{m=2}^N \chi_{nm} c_m (u_k - u_k^m)$$

$$\frac{\partial (\rho u_i)}{\partial t} + \partial_k \left(\sum_{n=1}^N \alpha_n \rho_n u_i^n u_k^n + \sum_{n=1}^N \alpha_n \rho_n - \alpha_1 \sigma_{ik} \right) = 0 \qquad \omega_j^n = e_{jkl} \frac{\partial w_k^n}{\partial x_l}$$

$$\frac{\partial \rho F_{ij}}{\partial t} + \partial_k (\rho F_{ij} u_k - \rho F_{kj} u_i) = 0$$

$$\frac{\partial}{\partial t} \rho \left(E + \frac{1}{2} u_i u_i \right) + \partial_k \left(\sum_{n=1}^N \alpha_n \rho_n u_k^n \left(e^n + \frac{1}{2} u_i^n u_i^n + \frac{1}{\rho} p_n \right) - u_i \alpha_1 \sigma_{ik} \right) = 0$$

Модель среды

Параметры состояния среды

- α_1 – объёмная доля твёрдой фазы (формации)
- α_n – объёмная доля жидкостей ($n=2, \dots, N$), $\alpha_1 + \sum_{n=2}^N \alpha_n = 1$
- ρ_n – массовая плотность фазы n
- $\rho = \sum_{n=1}^N \alpha_n \rho_n$ – полная массовая плотность среды
- $c_n = \frac{\alpha_n \rho_n}{\rho}$ – массовая доля фазы n , $c_1 + \sum_{n=2}^N c_n = 1$
- u_i^n – вектор скорости фазы n
- $w_i^n = u_i^1 - u_i^n$ – скорость фазы n относительно твёрдой фазы
- $u_i = \sum_{n=1}^N c_n u_i^n$ – скорость среды
- F_{ij} – градиент деформации среды
- S – энтропия среды
- $\phi = 1 - \alpha_1$ – пористость твёрдой фазы
- λ_{nm} – коэффициент релаксации
- χ_{nm} – коэффициент межфазного трения

Численный метод

- **Метод WENO-Рунге-Кутты 5-го порядка по пространству и времени**

- Система дифференциальных уравнений в двумерном случае может быть записана в следующей векторной форме:

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S(U)$$

- Разностная аппроксимация уравнения на прямоугольной сетке может быть записана в форме:

$$\frac{dU_{ij}}{dt} + \frac{F_{i+\frac{1}{2},j} - F_{i-\frac{1}{2},j}}{h_x} + \frac{G_{i,j+\frac{1}{2}} - G_{i,j-\frac{1}{2}}}{h_y} = S_{ij}$$

- Если значения потоков на гранях ячеек известны, можно получить дифференциальное уравнение для значений в ячейках сетки:

$$\frac{\partial U}{\partial t} = L(U) = - \frac{F_{i+\frac{1}{2},j} - F_{i-\frac{1}{2},j}}{h_x} - \frac{G_{i,j+\frac{1}{2}} - G_{i,j-\frac{1}{2}}}{h_y} + S_{ij}$$

- Для его решения используется метод Рунге-Кутты требуемого порядка точности.

Численный метод

- **Метод WENO-Рунге-Кутты 5-го порядка по пространству и времени**

- Для решения уравнения $\frac{\partial U}{\partial t} = L(U)$ используется метод Рунге-Кутты 5-го порядка точности

- $U^{(1)} = U^n + \frac{1}{2} \Delta t L(U^n)$

- $U^{(2)} = U^{(1)} + \frac{1}{2} \Delta t L(U^{(1)})$

- $U^{(3)} = U^{(2)} + \frac{1}{2} \Delta t L(U^{(2)})$

- $U^{(4)} = U^{(3)} + \frac{1}{2} \Delta t L(U^{(3)})$

- $U^{(5)} = U^{(4)} + \frac{1}{2} \Delta t L(U^{(4)})$

- $U^{n+1} = \frac{1}{9} U^n + \frac{2}{5} U^{(1)} + \frac{4}{9} U^{(3)} + \frac{2}{45} U^{(5)} + \Delta t L(U^{(5)})$

Численный метод

- **Метод WENO-Рунге-Кутты 5-го порядка по пространству и времени**

- Для определения значений потоков $F_{i\pm\frac{1}{2},j}$, $G_{i,j\pm\frac{1}{2}}$ на границах ячеек достаточно знать на них значения $U_{i\pm\frac{1}{2},j}$, $U_{i,j\pm\frac{1}{2}}$.
- Для вычисления потоков используется расщепление по пространству. На гранях ячеек используется аппроксимация WENO 5-го порядка с использованием значений из соседних ячеек.

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = 0 \qquad \frac{\partial U}{\partial t} + \frac{\partial G(U)}{\partial y} = 0$$

- Для вычисления потоков используется метод Лакса-Фридрихса:
 - По оси X: если U_L и U_R – значения слева и справа от точки $x_{i+\frac{1}{2}}$, то

$$F_{i+\frac{1}{2}}^{LF} = \frac{1}{2} (F(U_L) + F(U_R)) - \frac{1}{2} \frac{\Delta x}{\Delta t} (U_R - U_L)$$

- Аналогично по оси Y:

$$G_{i+\frac{1}{2}}^{LF} = \frac{1}{2} (G(U_L) + G(U_R)) - \frac{1}{2} \frac{\Delta y}{\Delta t} (U_R - U_L)$$

Алгоритм

Схема используемого метода Рунге-Кутты ($U^n \rightarrow U^{n+1}$):

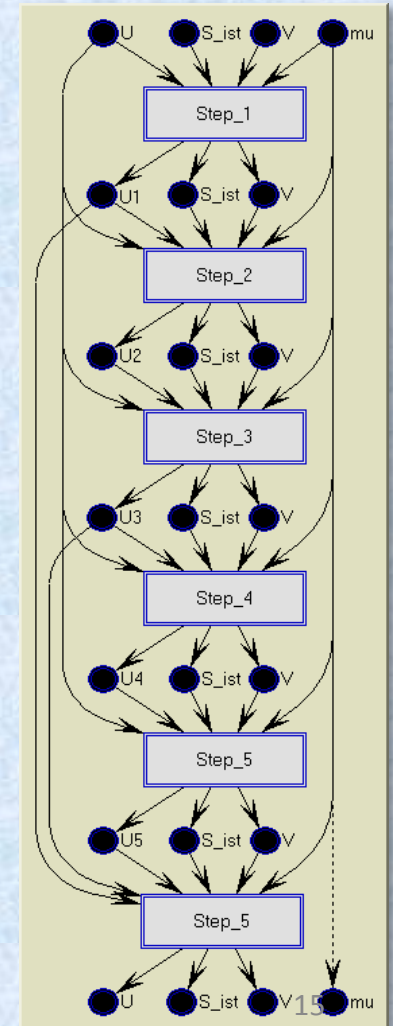
- $U_1 = U^n + \frac{1}{2} \tau L(U^n)$
- $U_2 = U_1 + \frac{1}{2} \tau L(U_1)$
- $U_3 = U_2 + \frac{1}{2} \tau L(U_2)$
- $U_4 = U_3 + \frac{1}{2} \tau L(U_3)$
- $U_5 = U_4 + \frac{1}{2} \tau L(U_4)$
- $U^{n+1} = \frac{1}{9} \cdot U + \frac{2}{5} \cdot U_1 + \frac{4}{9} \cdot U_3 + \frac{2}{45} \cdot U_5 + \frac{1}{45} \cdot \tau \cdot L(U_5)$

Вычисление оператора $L(U)$:

$$L(U) = -\frac{F_{i+\frac{1}{2},j} - F_{i-\frac{1}{2},j}}{h_x} - \frac{G_{i,j+\frac{1}{2}} - G_{i,j-\frac{1}{2}}}{h_y} + S_{i,j}$$

Вычисление потоков F и G : метод WENO 5-го порядка

- Используются только явные схемы



Алгоритм

Шаг метода Рунге-Кутты –
представление в виде графа
Обозначения:

– переменные –
характеристики среды
(2D массивы)

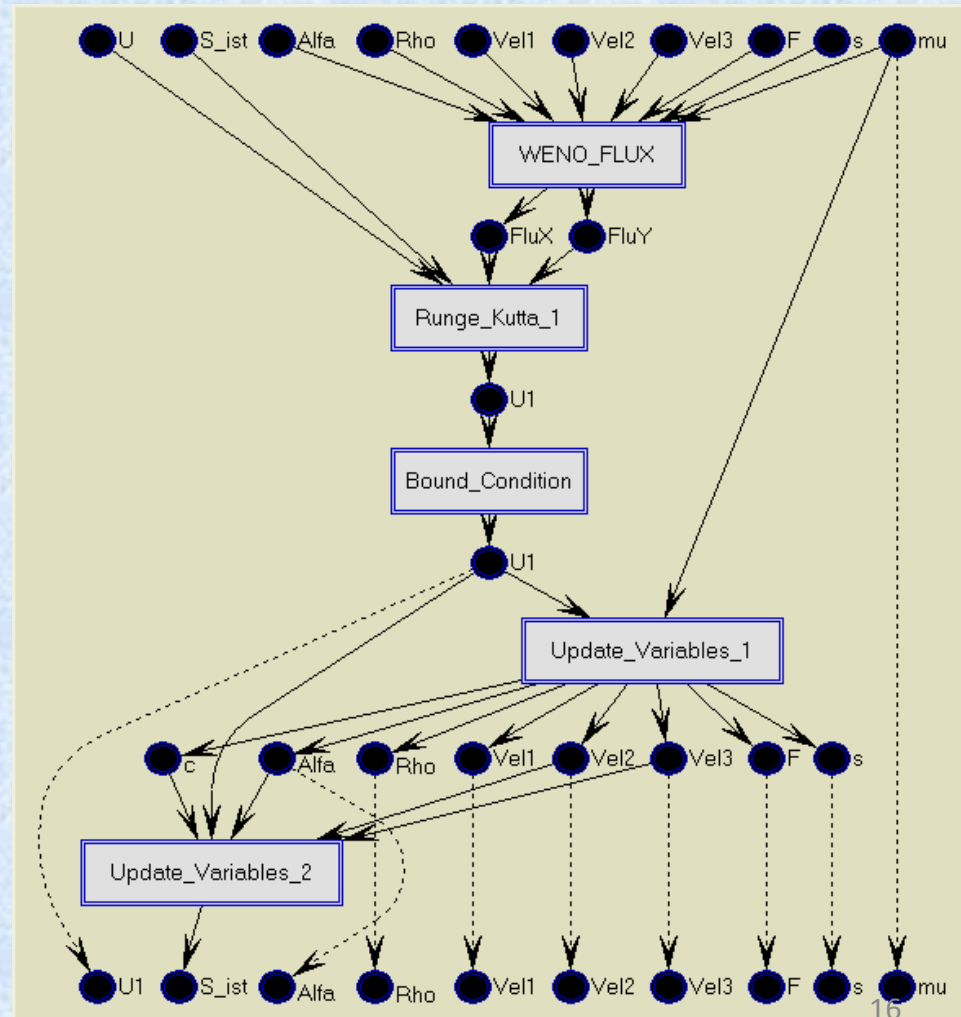
WENO_FLUX

– операции – вычисление
характеристик среды
(гнезда циклов – обход 2D
массива)

```
for (iy=ny1; iy<=ny2; iy++)  
  for (ix=nx1; ix<=nx2; ix++) { ... }
```

– зависимость по данным

– «идентичность»
переменных



Алгоритм

Шаг метода Рунге-Кутты –
представление в виде графа

Обозначения:

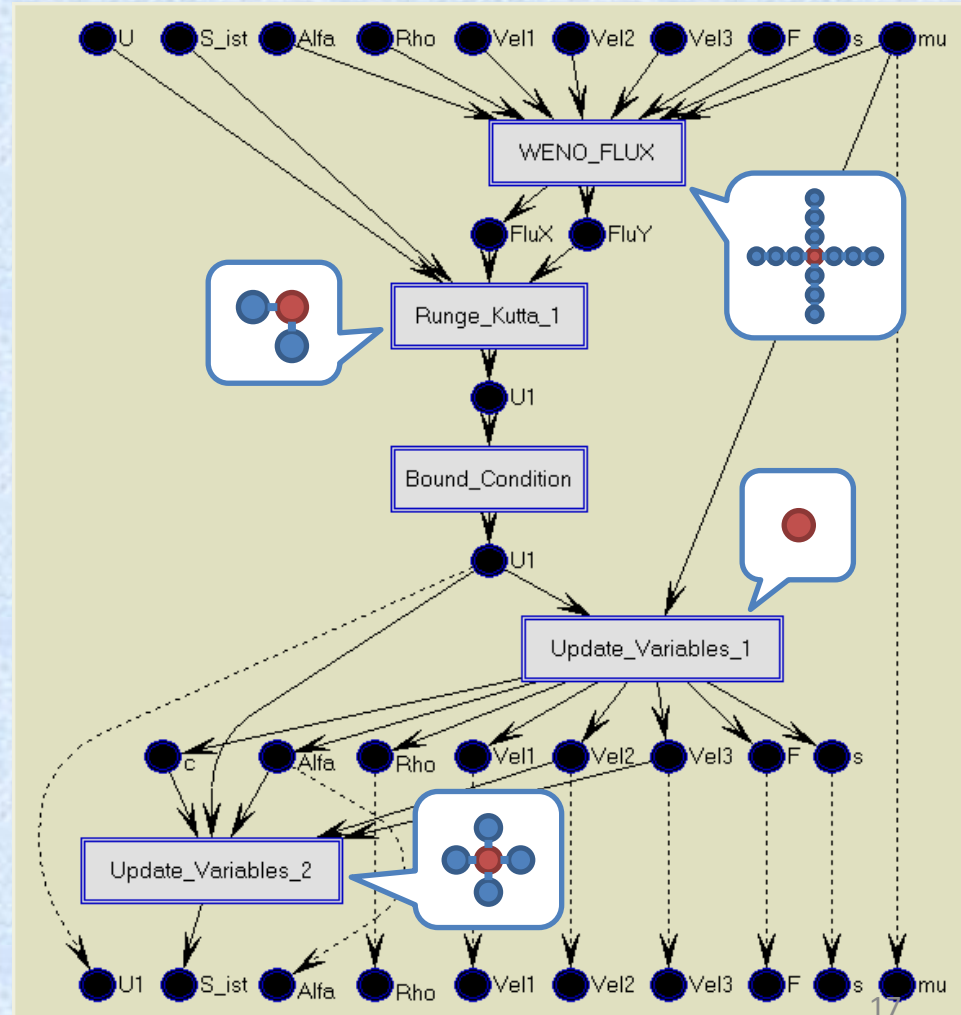
– переменные –
характеристики среды
(2D массивы)

– операции – вычисление
характеристик среды
(гнезда циклов – обход 2D
массива)

```
for (iy=ny1; iy<=ny2; iy++)
  for (ix=nx1; ix<=nx2; ix++) { ... }
```

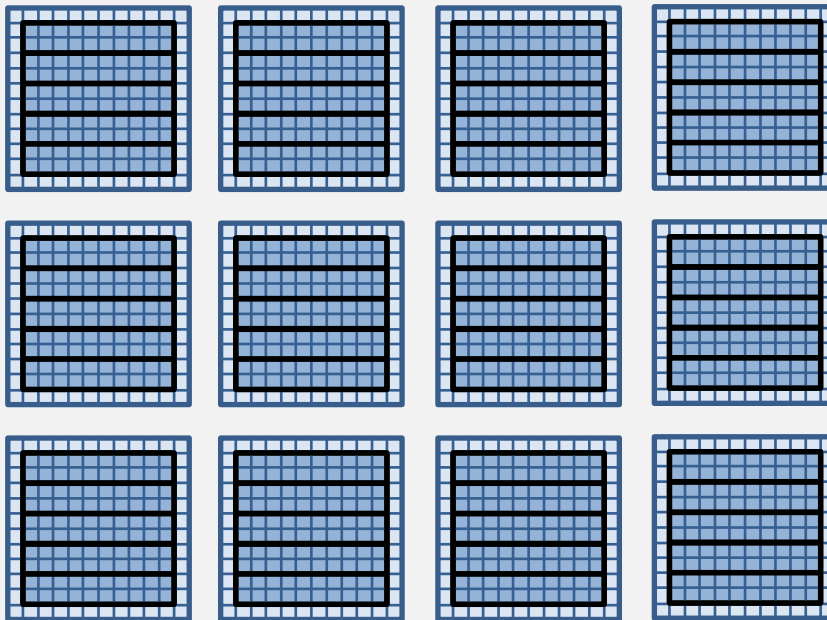
– зависимость по данным

– «идентичность»
переменных



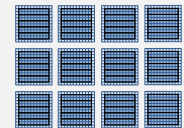
Параллельная реализация

- **Используется декомпозиция пространства моделирования**
 - Пространство моделирования делится на подобласти, которые распределяются по различным вычислителям
 - Между подобластями осуществляются обмены граничными значениями



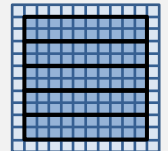
Параллельная реализация

- Распараллеливание для распределённой памяти
 - Используется MPI, 2D декомпозиция области



- Распараллеливание для общей памяти
 - Используется OpenMP, 1D декомпозиция области (по оси Y)

- Режим offload для использования ускорителей Intel Xeon Phi



- Используются средства компилятора Intel (#pragma offload...), 1D декомпозиция области (по оси Y)

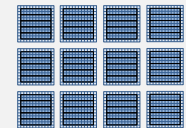
- Векторизация вычислений



- Выполнена средствами компилятора Intel (директивы компилятора #pragma... + небольшая модификация кода)

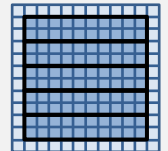
Параллельная реализация

- Распараллеливание для распределённой памяти
 - Используется MPI, 2D декомпозиция области



- Распараллеливание для общей памяти
 - Используется OpenMP, 1D декомпозиция области (по оси Y)

- Режим offload для использования ускорителей Intel Xeon Phi



- Используются средства компилятора Intel (#pragma offload...), 1D декомпозиция области (по оси Y)

- Векторизация вычислений



- Выполнена средствами компилятора Intel (директивы компилятора #pragma... + небольшая модификация кода)

Параллельная реализация

- Параметры запуска, задаваемые через систему очередей кластера
 - Количество узлов кластера
 - Количество MPI-процессов на хост-процессоры узла
 - Количество MPI-процессов на ускоритель KNC (для режима Native)
- Параметры запуска, задаваемые через MPI
 - Количество MPI-процессов
- Параметры запуска программного комплекса
 - Размеры 2D решётки процессов
 - Количество потоков на процесс, отдельно для хост-процессоров и ускорителей KNC
 - Распределение работ между хост-процессорами и ускорителями KNC для режима offload

План доклада

- **Описание задачи**
 - Введение
 - Математическая модель
 - Численный метод
 - Алгоритм
 - Параллельная реализация
- **Описание кластеров**
 - Архитектура
 - Вопросы использования
- **Результаты сравнительного тестирования**

Сравнение кластеров

МВС-10П (МСЦ РАН)

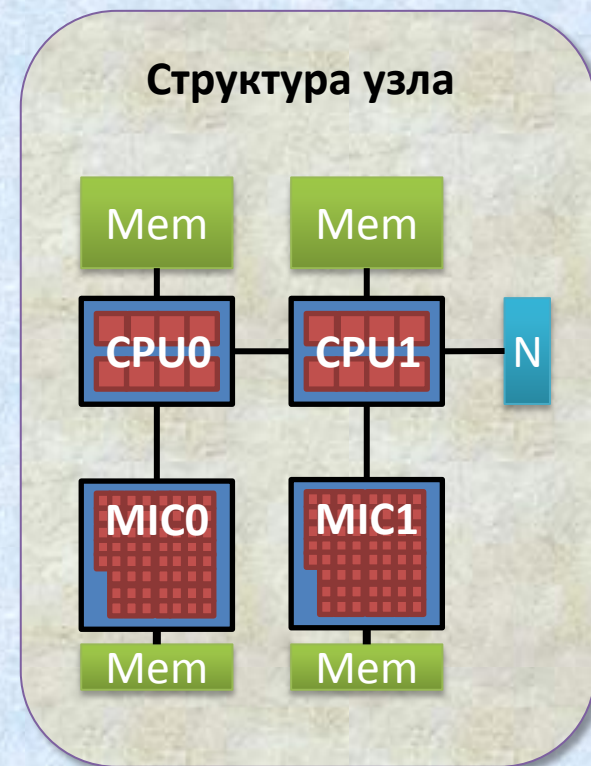
- «Основной» раздел (2013 год) – ресурсы узла:
 - 2 × процессора Intel Xeon (**Sandy Bridge**)
 - 2 × ускорителя Intel Xeon Phi (**KNC** – Knight's Corner)
- Раздел PetaStream...
- Раздел Broadwell...
- Раздел KNL...

НКС-1П (ССКЦ СО РАН)

- Раздел Broadwell (2017 год) – ресурсы узла:
 - 2 × процессора Intel Xeon (**Broadwell**)
- Раздел KNL (2017 год) – ресурсы узла:
 - 1 × процессор Intel Xeon Phi (**KNL** – Knight's Landing)

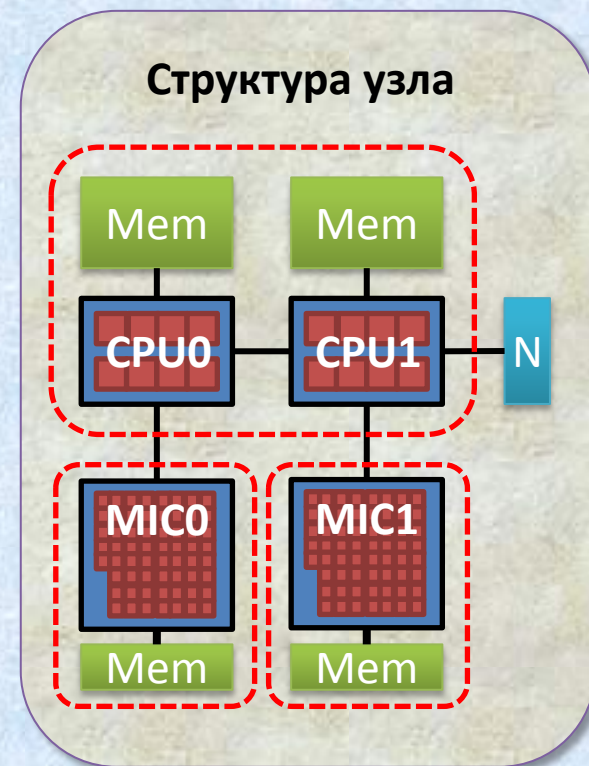
Кластер МВС-10П

- Узел кластера:
 - хост-система
 - 2 × Xeon E5-2690, 2.9 ГГц (Sandy Bridge), 8 ядер × 2 потока
 - 32 аппаратных потока на узел
 - 64 Гб памяти на узел
 - **185.6 GFLOPS** пиковая произв-ть
 - два ускорителя Intel Xeon Phi
 - Xeon Phi 7110X (KNC), 1.094 ГГц, 61 ядро × 4 потока
 - 244 аппаратных потока на ускоритель
 - 8 Гб памяти на ускоритель
 - **2 × 1067.7 GFLOPS** пиковая произв-ть



Кластер МВС-10П

- Узел кластера:
 - хост-система
 - 2 × Xeon E5-2690, 2.9 ГГц (Sandy Bridge), 8 ядер × 2 потока
 - 32 аппаратных потока на узел
 - 64 Гб памяти на узел
 - **185.6 GFLOPS** пиковая произв-ть
 - два ускорителя Intel Xeon Phi
 - Xeon Phi 7110X (KNC), 1.094 ГГц, 61 ядро × 4 потока
 - 244 аппаратных потока на ускоритель
 - 8 Гб памяти на ускоритель
 - **2 × 1067.7 GFLOPS** пиковая произв-ть



Кластер НКС-1П

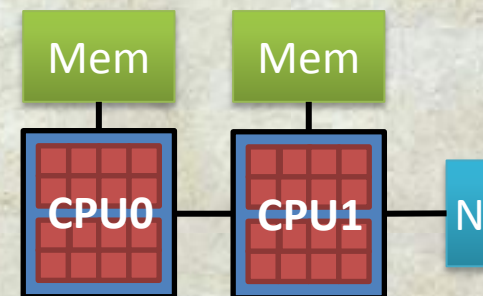
- Узел кластера (раздел Broadwell):

- 2 × Xeon E5-2697A v4, 2.6 (3.6) ГГц (Broadwell), 16 ядер × 2 потока
- 32 аппаратных потока на узел
- 128 ГБ памяти на узел
- **1331.2 GFLOPS** пиковая произв-ть

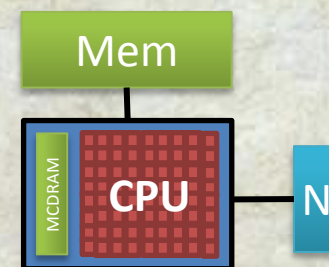
- Узел кластера (раздел KNL):

- 1 × Xeon Phi 7290, 1.5 (1.7) ГГц (KNL), 72 ядра × 4 потока
- 288 аппаратных потока на узел
- 96 ГБ памяти на узел
- 16 GB памяти MCDRAM (кэш)
- **3456 GFLOPS** пиковая произв-ть

Структура узла



Структура узла



Сравнение процессоров

- Пиковая производительность узла

	FLOPs на команду	Команд за такт	Длина вектора	GHz	Ядер в проц.	Проц-ров в узле	GFLOPS на узел
MBC-10П							
Sandy Bridge	1	× 1	× 4 (AVX)	× 2.9	× 8	× 2	= 185.6
KNC	2 (FMA)	× 1	× 8 (IMCI)	× 1.1	× 61	× 2	= 2135.5
HKC-1П							
Broadwell	2 (FMA)	× 2	× 4 (AVX2)	× 2.6	× 16	× 2	= 1331.2
KNL	2 (FMA)	× 2	× 8 (AVX-512)	× 1.5	× 72	× 1	= 3456

Сравнение процессоров

- Подсистема памяти

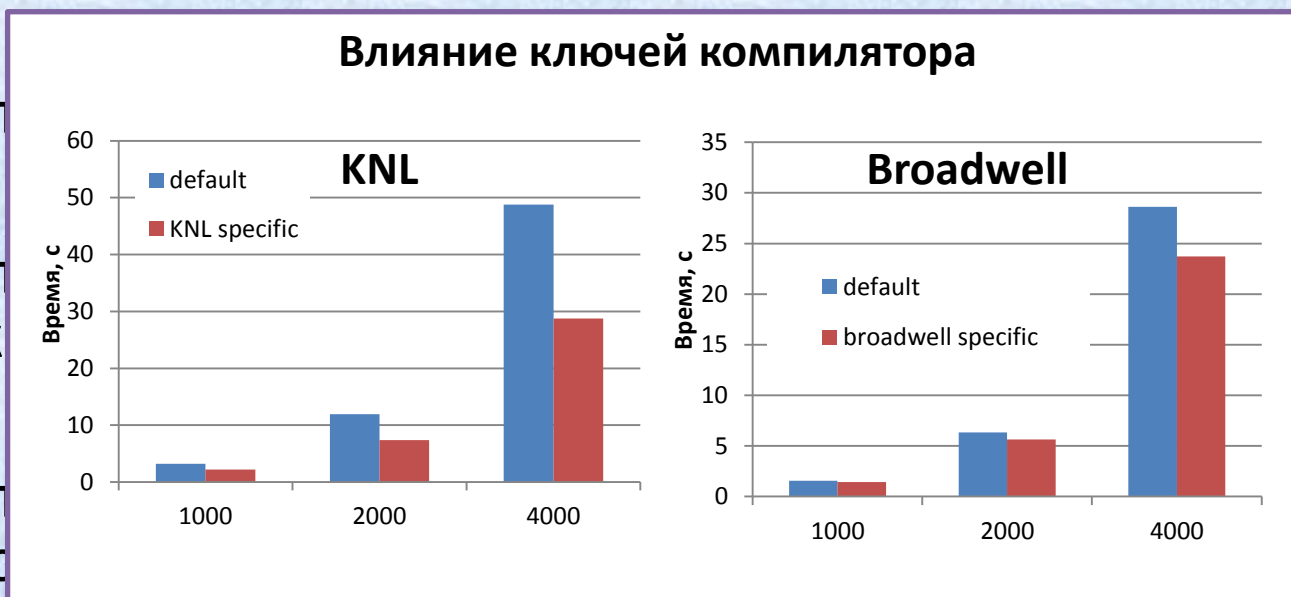
	Кэш L1	Кэш L2	Кэш L3	Локальная память
МВС-10П				
Sandy Bridge	8 × 32 KB	8 × 256 KB	20 MB	-
KNC	61 × 32 KB	61 × 512 KB	-	8 GB GDDR5 352 GB/s
НКС-1П				
Broadwell	16 × 32 KB	16 × 256 KB	2 × 20 MB	-
KNL	72 × 32 KB	36 × 1 MB	-	16 GB MCDRAM 400+ GB/s

Вопросы оптимального использования ресурсов

- Правильное использование компилятора
- Оптимальный способ 2D декомпозиции области
- Оптимальное соотношение процессов и потоков в узле / ускорителе
- Оптимальное соотношение работы между хост-процессорами и ускорителями KNC в узлах кластера МВС-10П
- Использование ускорителей: режимы Native и Offload

Вопросы оптимального использования ресурсов

- Правильное использование компилятора



- Оп

- Оп уск

- Оп прс

МВС-10П

в узле /

ера

- Использование ускорителей: режимы Native и Offload

Вопросы оптимального использования ресурсов

- Правильное использование компилятора
- Оптимальный способ 2D декомпозиции области
- Оптимальное соотношение процессов и потоков в узле / ускорителе
- Оптимальное соотношение работы между хост-процессорами и ускорителями KNC в узлах кластера МВС-10П
- Использование ускорителей: режимы Native и Offload

Вопросы оптимального использования ресурсов

- Правильное использование компилятора

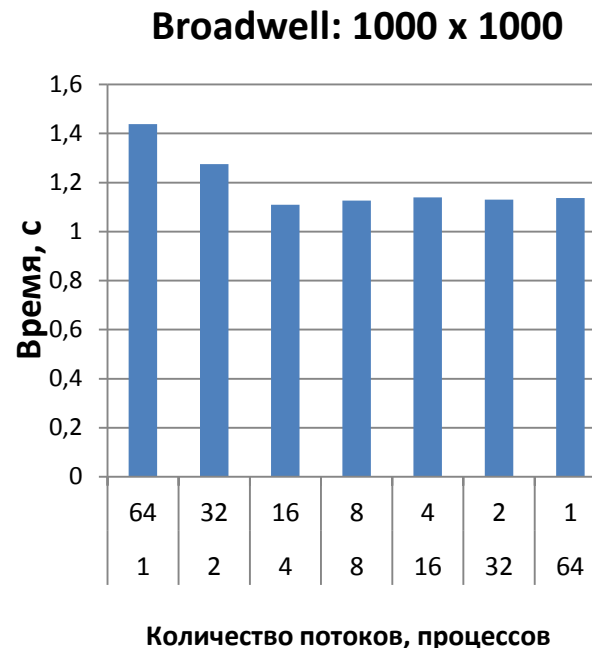
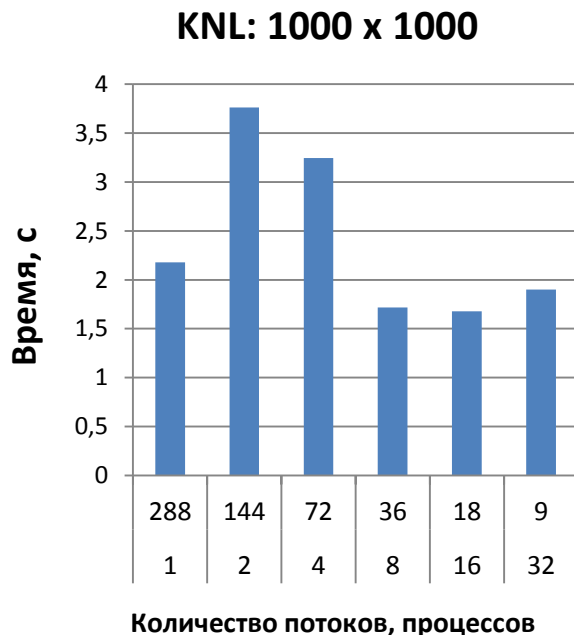
- Опт

- Опт
уск

- Опт
пр
МВ

- Исп

Влияние соотношения процессов и потоков в узле



в узле /

ера

flood

Вопросы оптимального использования ресурсов

- Правильное использование компилятора
- Оптимальный способ 2D декомпозиции области
- Оптимальное соотношение процессов и потоков в узле / ускорителе
- Оптимальное соотношение работы между хост-процессорами и ускорителями KNC в узлах кластера МВС-10П
- Использование ускорителей: режимы Native и Offload

Вопросы оптимального использования ресурсов

- Правильное использование компилятора

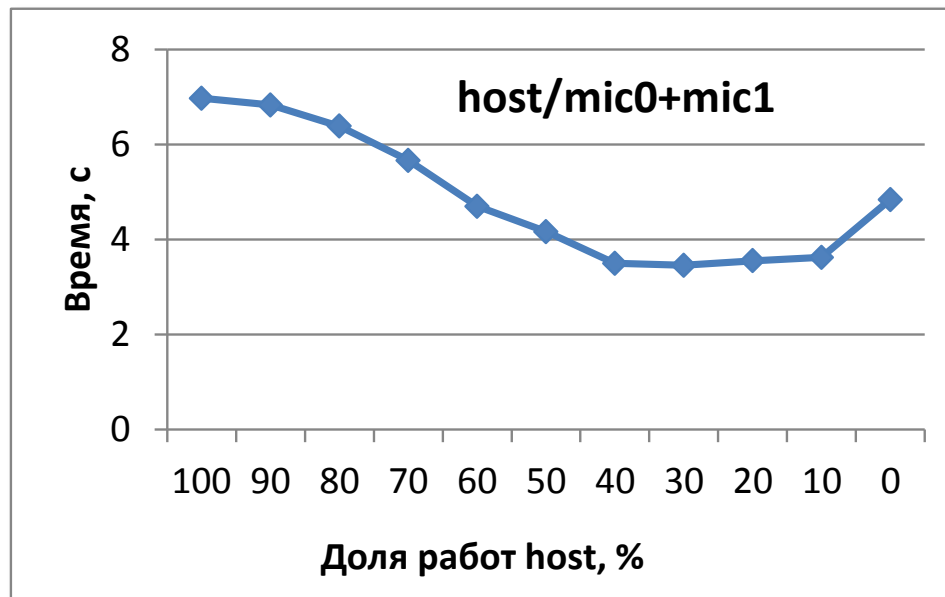
- Опт

- Опт
уск

- Опт
прс
МВ

- Исп

Влияние соотношения работ между хост-ядрами и ускорителями



в узле /

ера

ifload

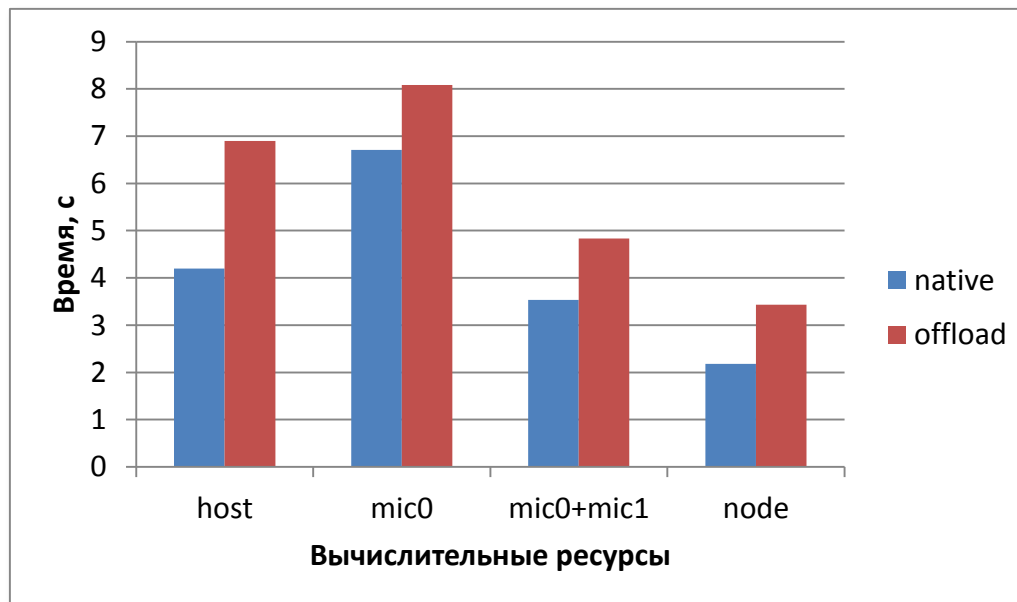
Вопросы оптимального использования ресурсов

- Правильное использование компилятора
- Оптимальный способ 2D декомпозиции области
- Оптимальное соотношение процессов и потоков в узле / ускорителе
- Оптимальное соотношение работы между хост-процессорами и ускорителями KNC в узлах кластера МВС-10П
- Использование ускорителей: режимы Native и Offload

Вопросы оптимального использования ресурсов

- Правильное использование компилятора

Сравнение режимов:
режим Offload на данной задаче менее эффективен



- Опт

- Опт
уск

- Опт
прс
МВ

- Исп

в узле /

ера

ifload

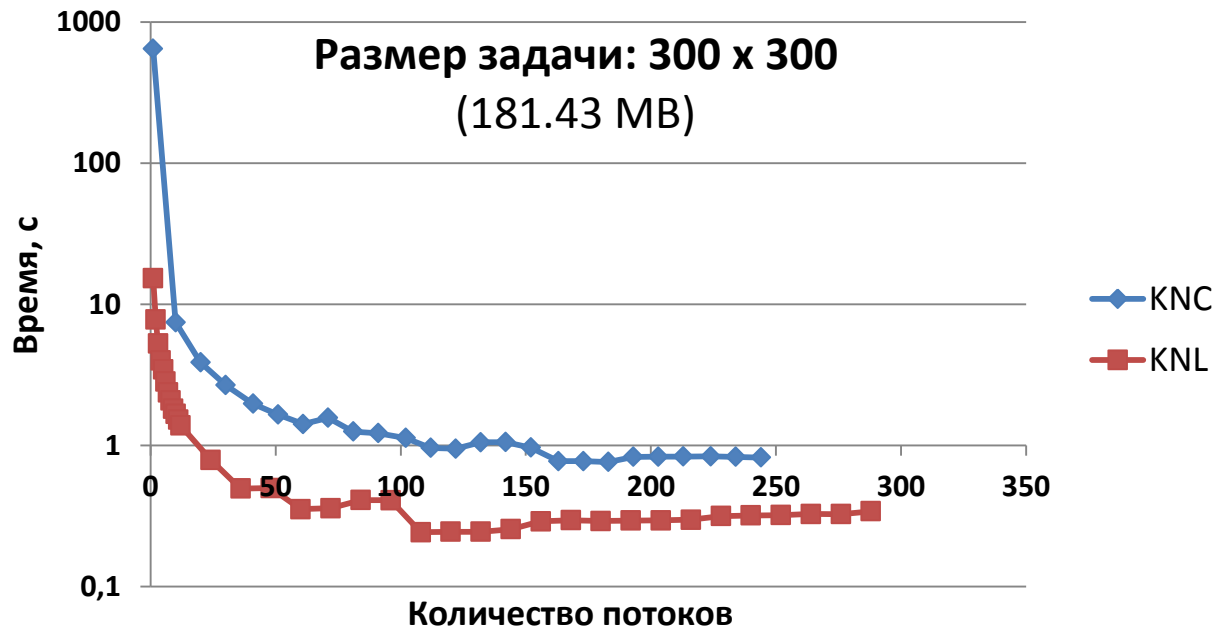
Вопросы оптимального использования ресурсов

- Правильное использование компилятора
- Оптимальный способ 2D декомпозиции области
- Оптимальное соотношение процессов и потоков в узле / ускорителе
- Оптимальное соотношение работы между хост-процессорами и ускорителями KNC в узлах кластера МВС-10П
- Использование ускорителей: режимы Native и Offload

План доклада

- **Описание задачи**
 - Введение
 - Математическая модель
 - Численный метод
 - Алгоритм
 - Параллельная реализация
- **Описание кластеров**
 - Архитектура
 - Вопросы использования
- **Результаты сравнительного тестирования**

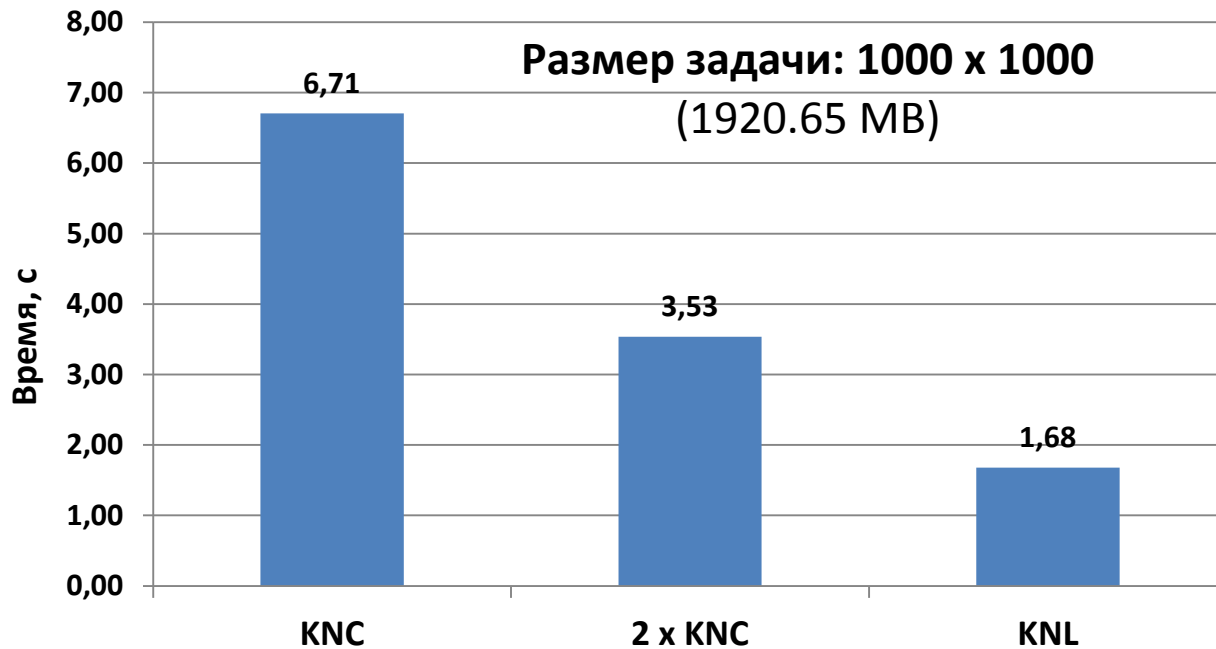
Сравнение KNC и KNL



Приведено время
счёта одного шага
по времени

- Время счёта на одном ядре:
 - KNC: **644.4 с**
 - KNL: **15.3 с** - быстрее в 42 раза

Сравнение KNC и KNL



Расчёт на всех ядрах

Приведено время счёта одного шага по времени

Время счёта на всех ядрах:

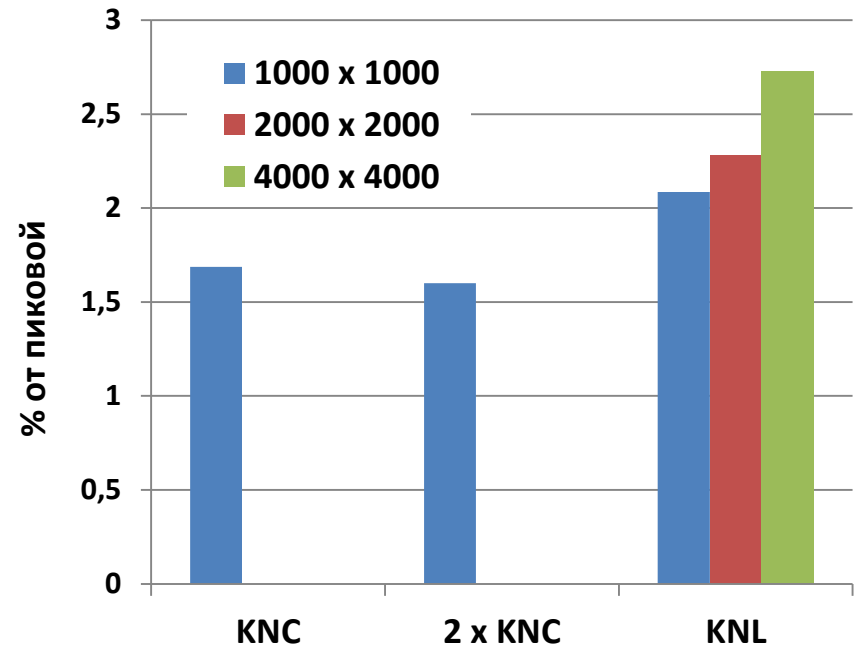
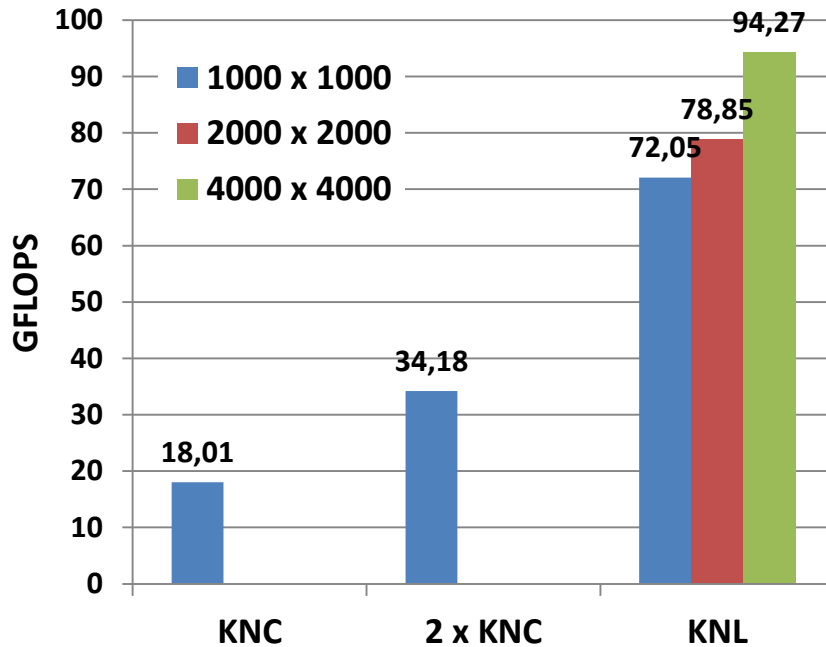
- KNC: 6.71 с
- KNL: 1.68 с - **быстрее в 4 раза**

Время счёта на всех ядрах узла:

- 2 x KNC: 3.53 с
- KNL: 1.68 с - **быстрее в 1.9 раза**

Сравнение KNC и KNL

Достигнутая производительность

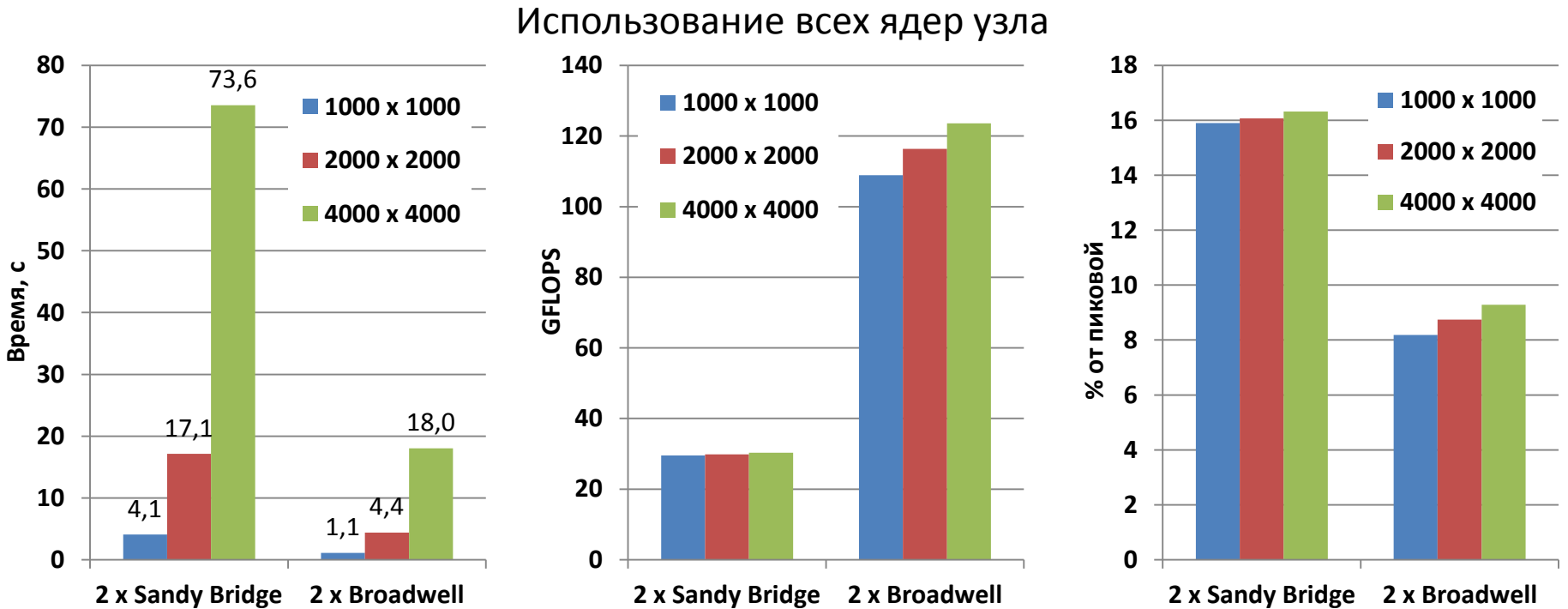


1000 x 1000 : **1 920.65 MB**

2000 x 2000 : **7 625.38 MB**

4000 x 4000 : **30 387.38 MB**

Сравнение Sandy Bridge и Broadwell



Ускорение на 2 x Broadwell по сравнению с 2 x Sandy Bridge:

- 1000 x 1000 : быстрее в **3.7** раза
- 2000 x 2000 : быстрее в **3.9** раза
- 4000 x 4000 : быстрее в **4.1** раза

Размеры задач:

1000 x 1000 : **1 920.65** MB

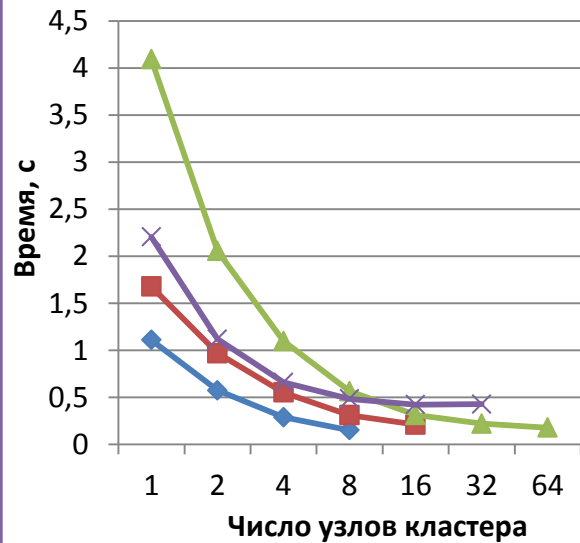
2000 x 2000 : **7 625.38** MB

4000 x 4000 : **30 387.38** MB

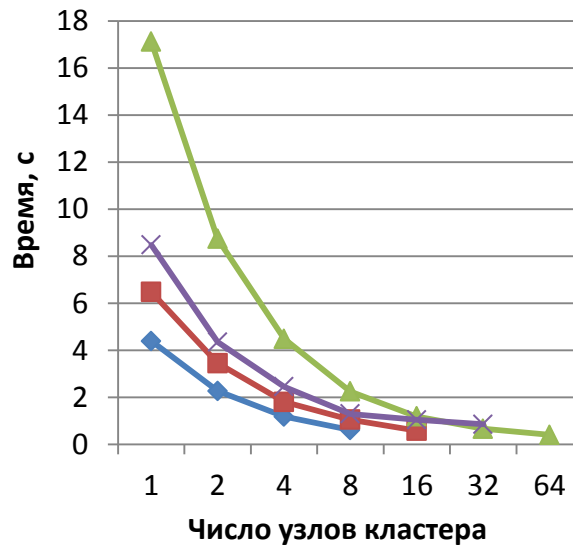
Сравнение МВС-10П и НКС-1П

Время счёта задачи

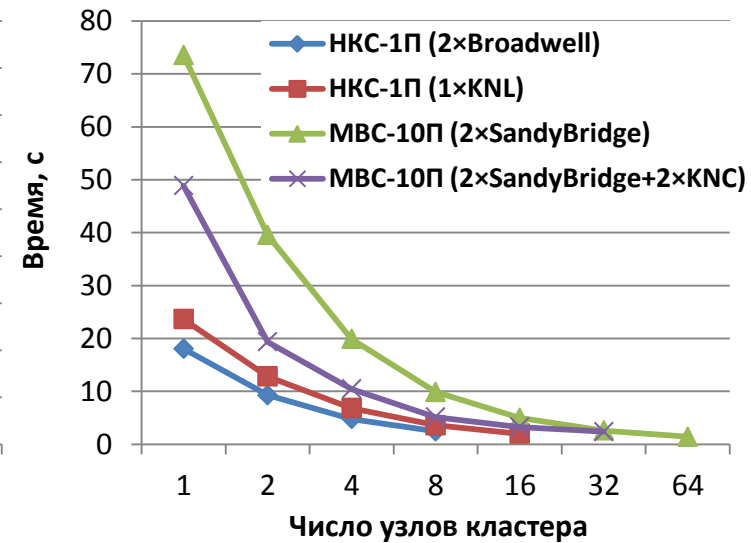
1000×1000



2000×2000



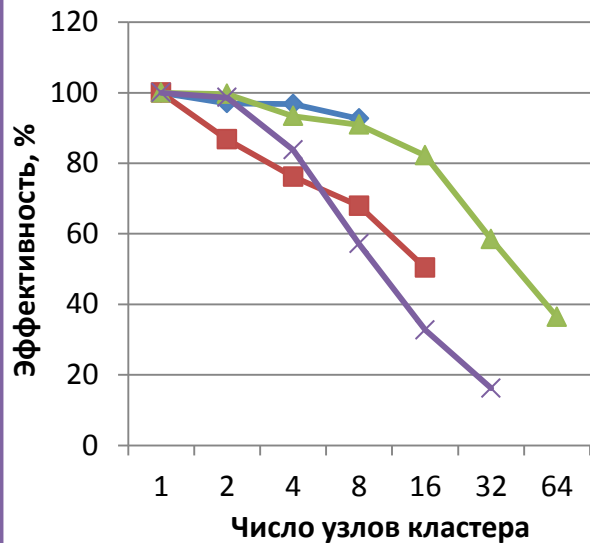
4000×4000



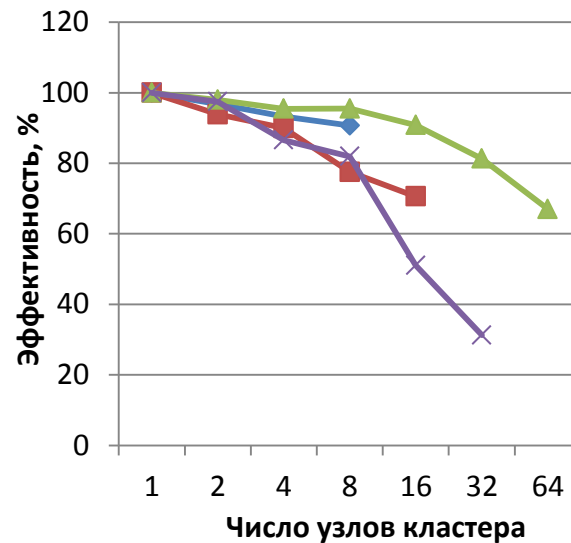
Сравнение МВС-10П и НКС-1П

Эффективность распараллеливания по узлам кластера

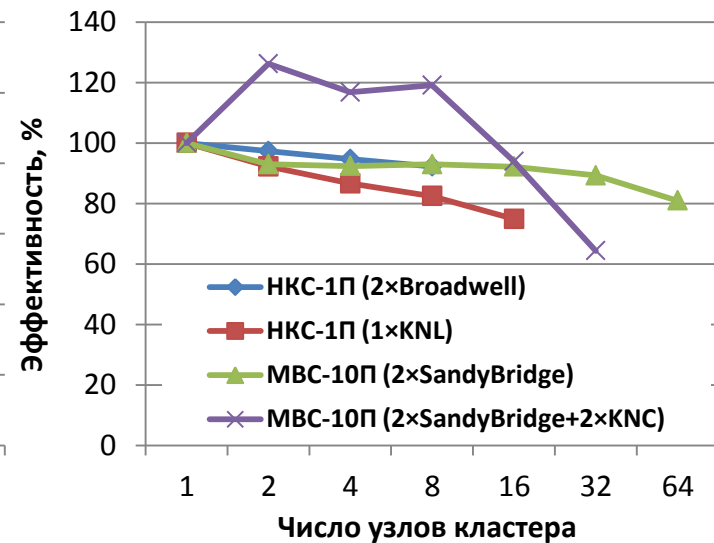
1000×1000



2000×2000

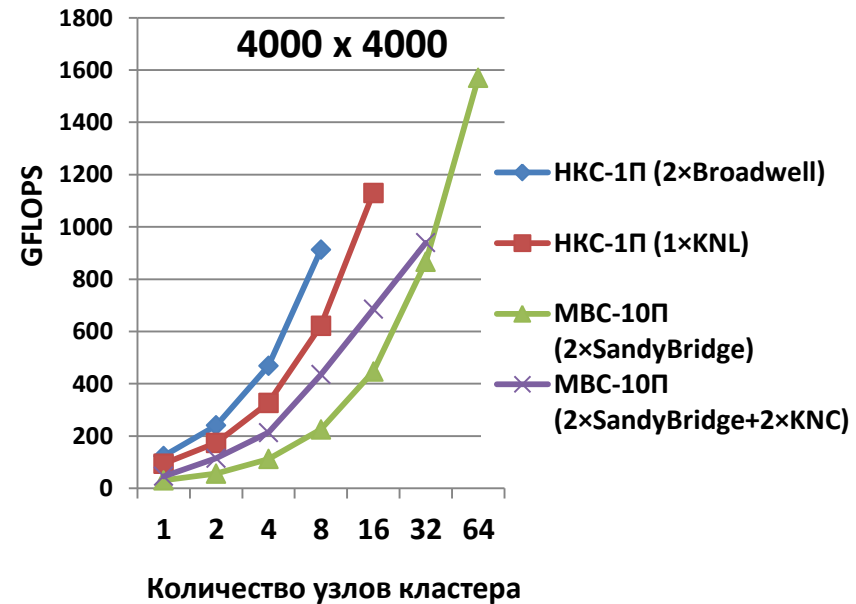
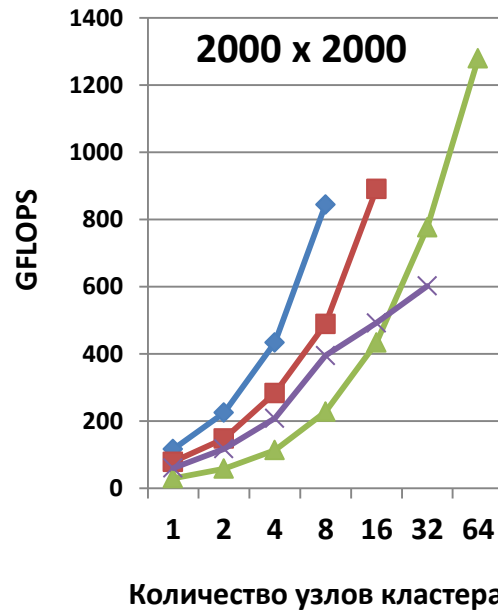
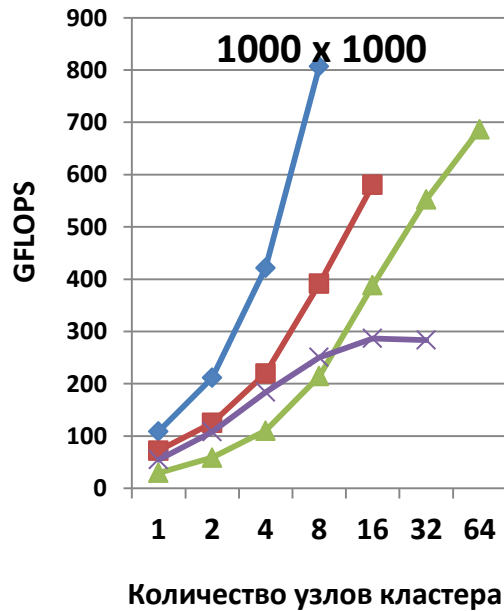


4000×4000



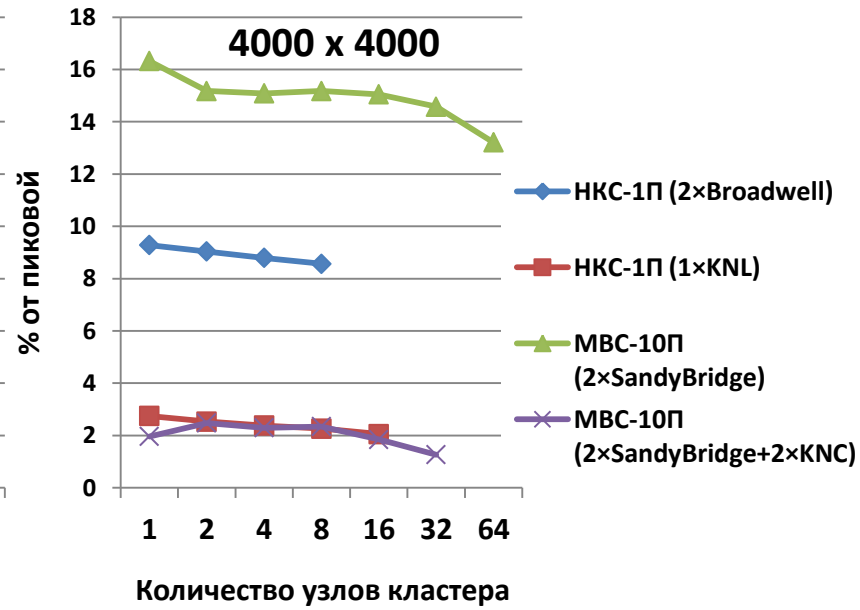
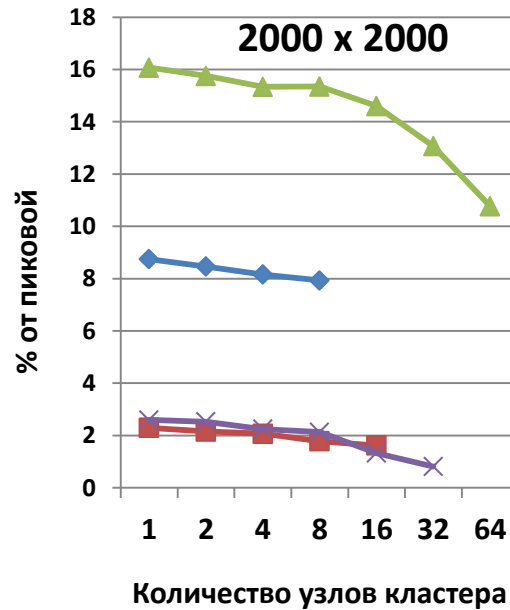
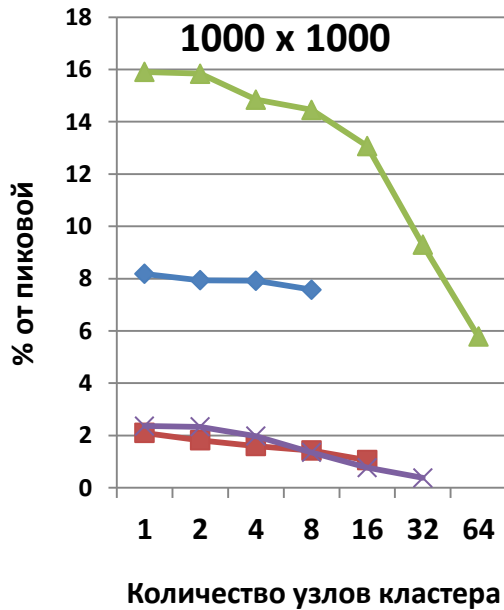
Сравнение МВС-10П и НКС-1П

Достигнутая производительность



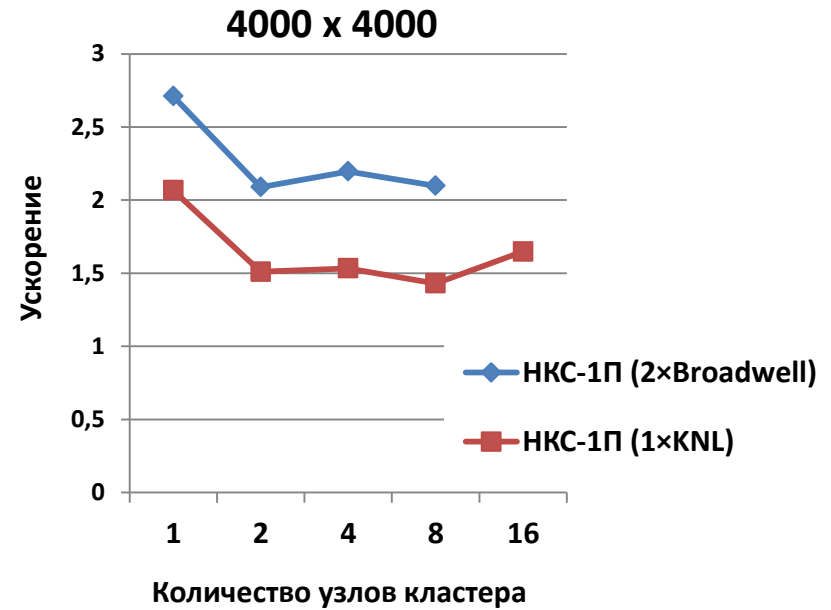
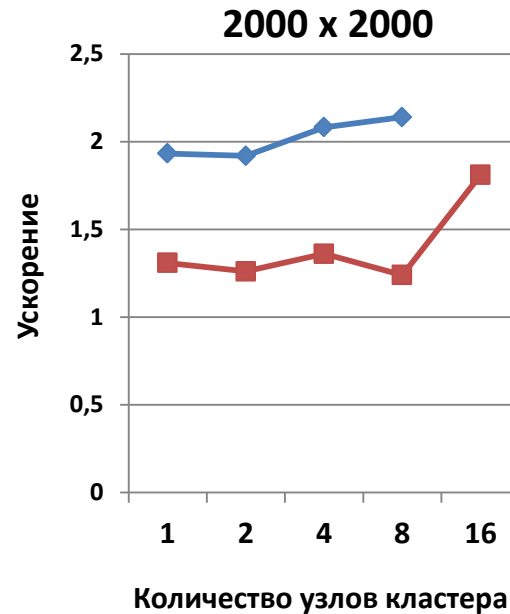
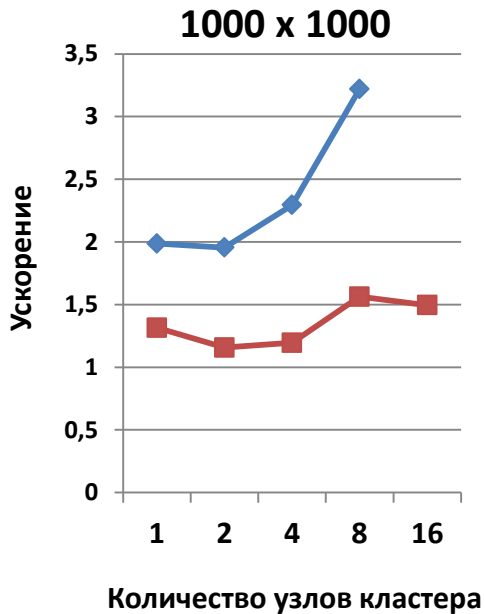
Сравнение МВС-10П и НКС-1П

Достигнутая производительность, процент от пиковой



Сравнение МВС-10П и НКС-1П

Ускорение при использовании НКС-1П по сравнению с МВС-10П



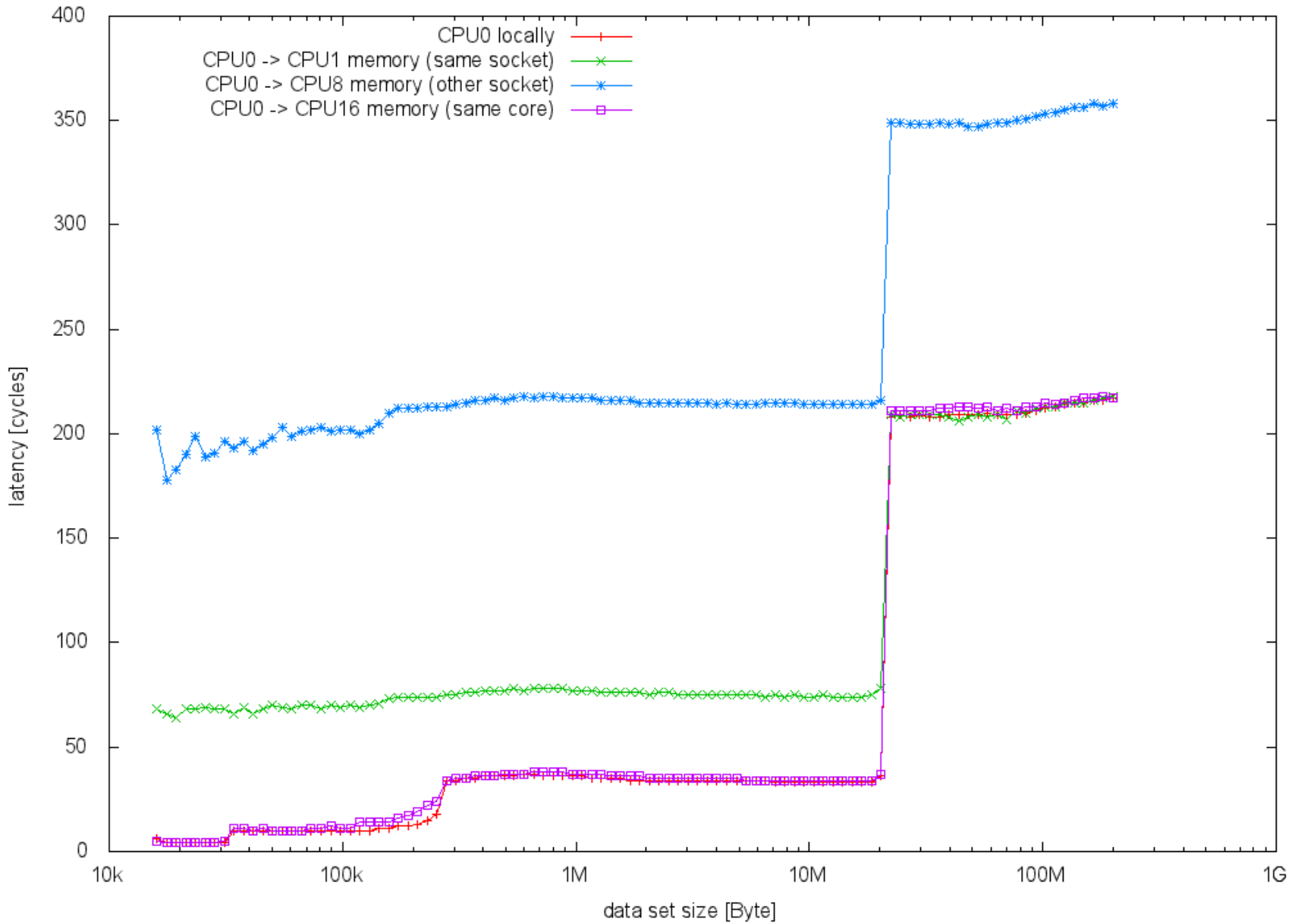
Выводы

- Эффективное использование отдельных разделов кластера НКС-1П проще, чем кластера МВС-10П, благодаря однородной структуре.
- Использование кластера НКС-1П позволяет ускорить расчёты по представленной программе по сравнению с кластером МВС-10П более чем в 1.5 раза на том же числе узлов.

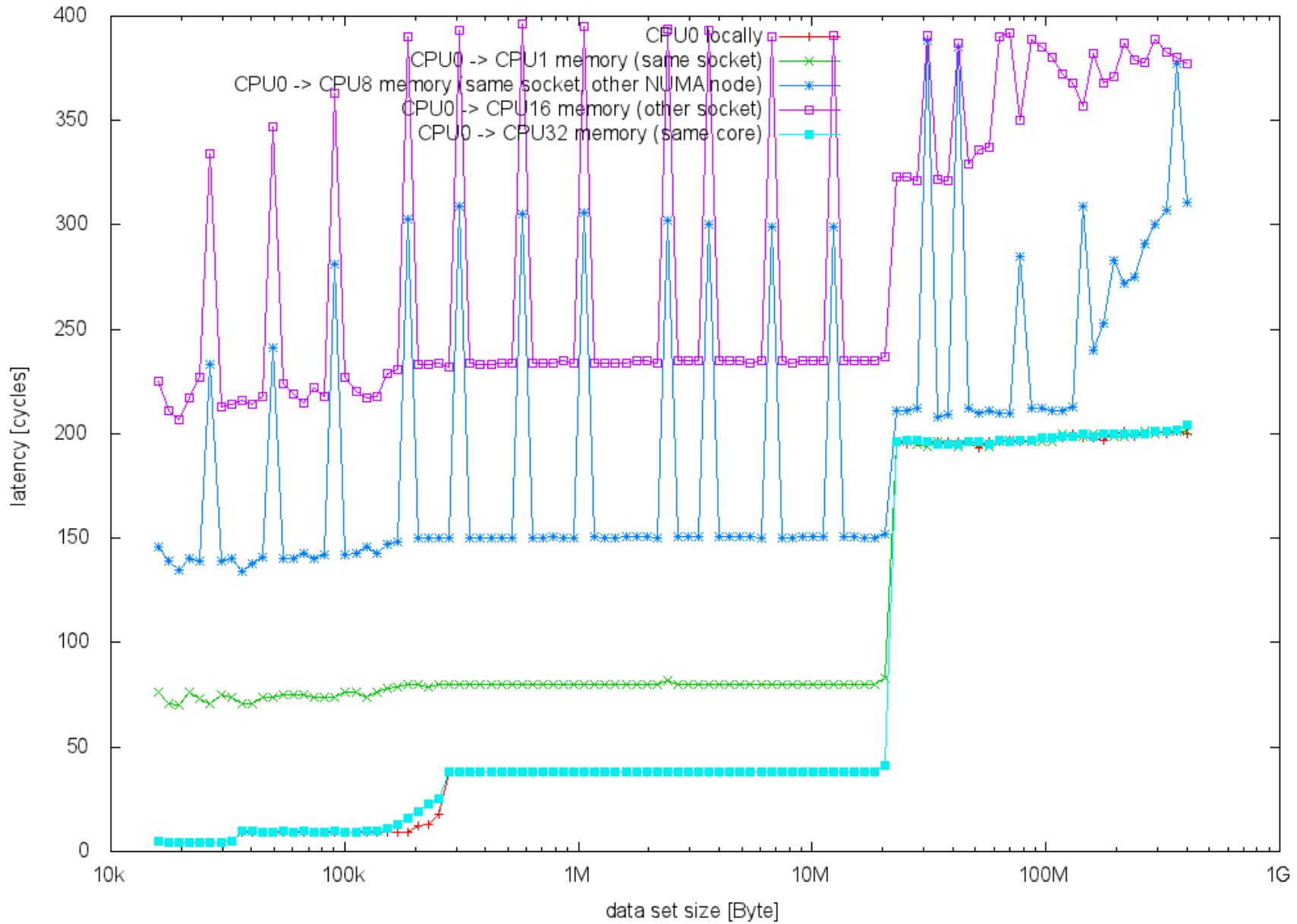
Спасибо за внимание!

Дополнительные слайды

MVS-10P Sandy Bridge: read latency



NKS-1P Broadwell : READ latency



NKS-1P KNL : read latency

