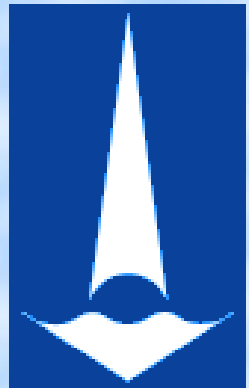


**Об особенностях решения
больших систем линейных
алгебраических уравнений на
многопроцессорных
вычислительных системах с
различной архитектурой**

Б.И. Краснопольский

НИИ механики МГУ

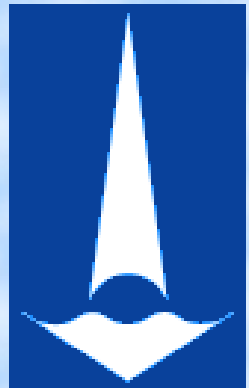


Параллельные Вычислительные Технологии (ПаВТ'2011)
29 - 31 марта 2011 г., Москва

О *некоторых* особенностях
~~**Об особенностях**~~ решения
больших систем линейных
алгебраических уравнений на
многопроцессорных
вычислительных системах с
различной архитектурой

Б.И. Краснопольский

НИИ механики МГУ



Параллельные Вычислительные Технологии (ПаВТ'2011)
29 - 31 марта 2011 г., Москва

Мотивация работы

Разработка параллельной программы для моделирования трехмерных гидродинамических течений вязкой несжимаемой теплопроводной жидкости для произвольных областей на неравномерных прямоугольных сетках

Вычислительная задача

Наиболее трудоемкая задача с точки зрения вычислений - решение системы линейных алгебраических уравнений вида

$$Ax = b$$

Переупорядоченный BiCGStab

Основные операции:

$$y = Ax, \quad y = K^{-1}x, \quad z = \alpha x + \beta y, \quad \alpha = (x, y)$$

Classical BiCGStab

1. $t = Av$
2. $\alpha = (v, x)$
3. $s = y + \alpha z$

Reordered BiCGStab

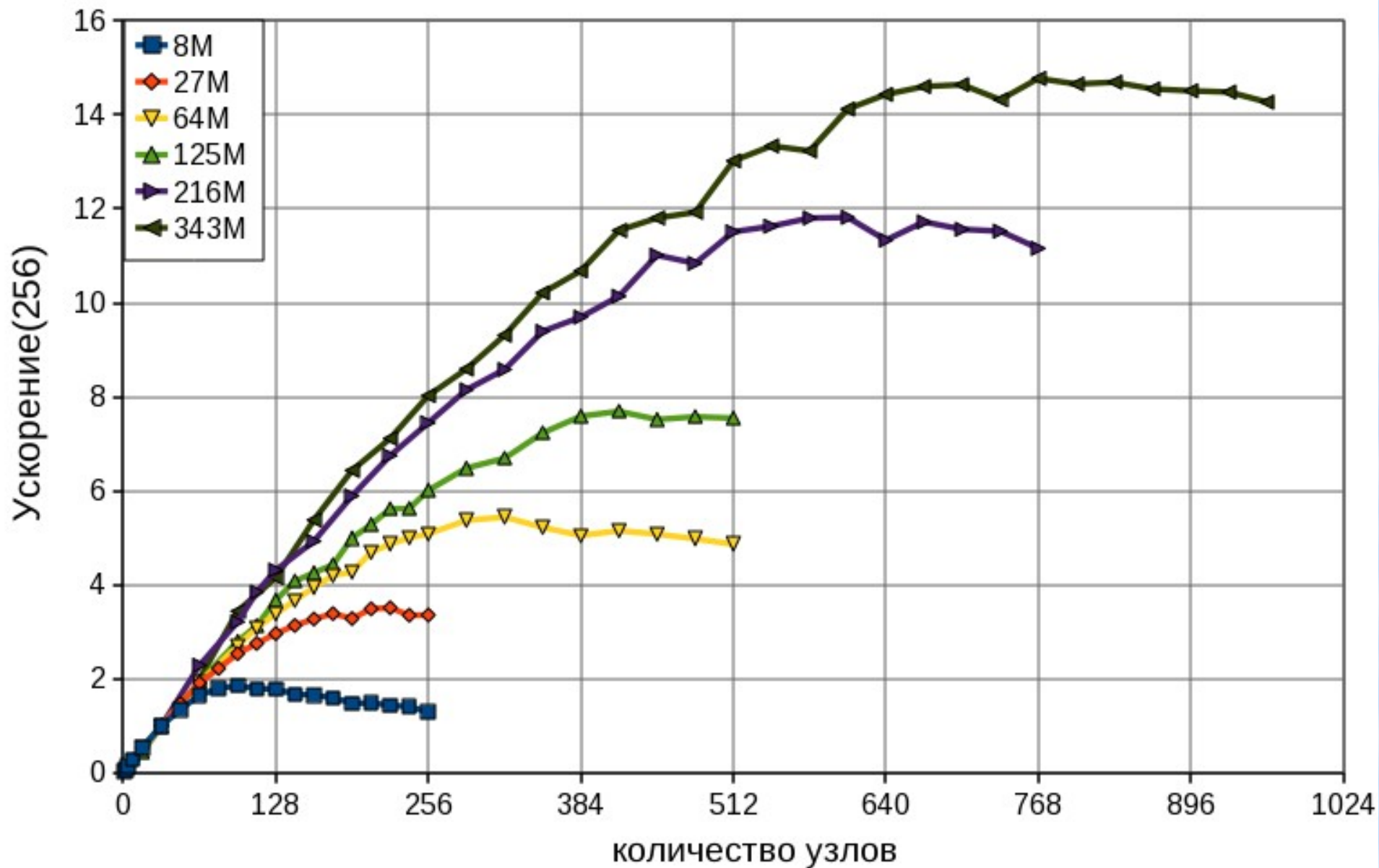
1. $t = Av$
2. $\alpha = (v, x)$
3. $p = K^{-1}r$
4. $s = y + \alpha z$

Особенности реализации вычислительных методов

Библиотека *SparseLinSol*:

- Гибридная модель **MPI+Posix Shared Memory**
- Для работы с данными использованы **CRS** и **DMSR** форматы хранения разреженных матриц
- При построении предобуславливания использована *open-source* библиотека ***hypre***
- Для разбиения исходной матрицы по процессам использована библиотека ***Pt-Scotch*** (исследованы библиотеки ***METIS/ParMETIS, hMETIS, Scotch, PaToH***)

Результаты предыдущих тестов



Тестовая задача

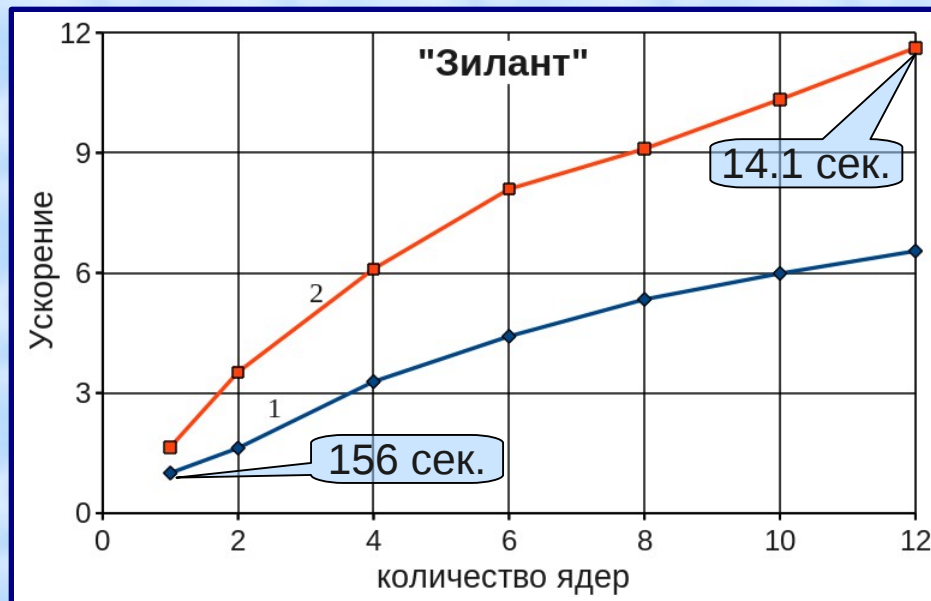
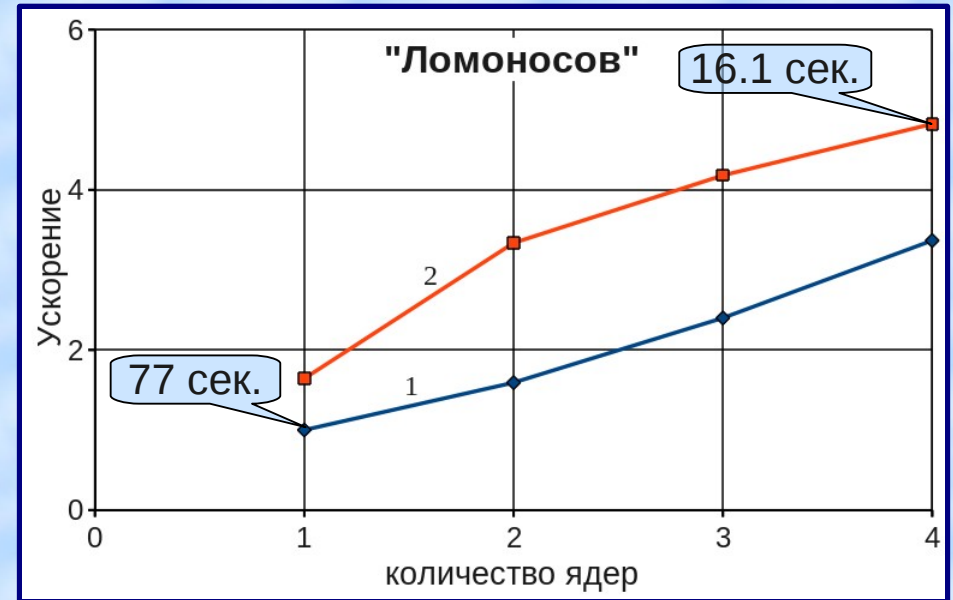
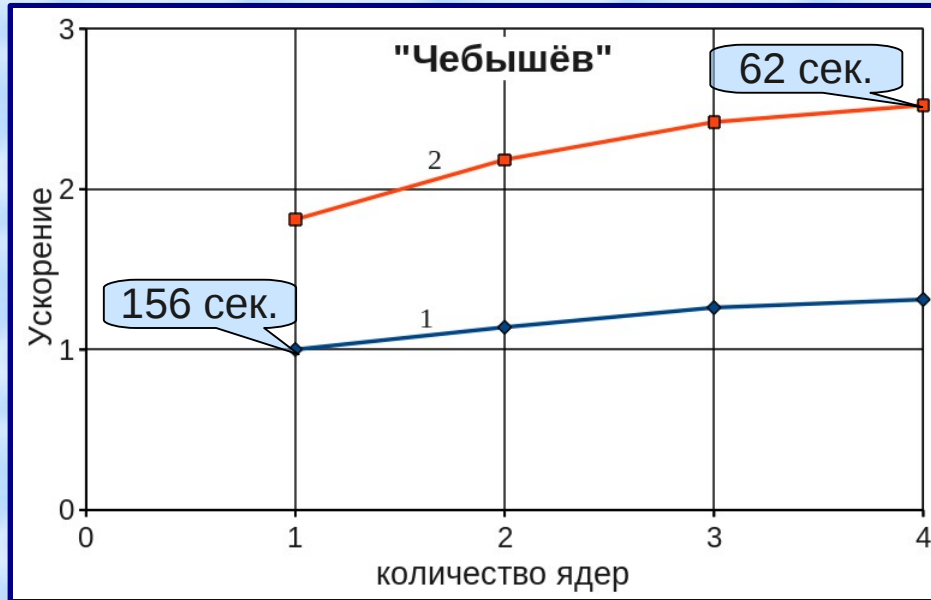
- В качестве тестовой задачи была выбрана задача Неймана для уравнения Пуассона с постоянной правой частью в кубической области на равномерной сетке (7-точечный шаблон пространственной аппроксимации)
- Размер матрицы системы 8 млн. неизвестных
- Фиксированное количество итераций метода $N_{it} = 20$
- Ускорение решения задачи определено как:

$$P_N = \frac{T_1}{T_N}$$

Тестовые платформы

	“Чебышёв”	“Ломоносов”	“Зилант”
Тип и модель процессора	Intel Harpertown, E5472	Intel Nehalem, X5570	AMD Magny Cours, Opteron 6174
Тактовая частота	3.0 ГГц	2.93 ГГц	2.2 ГГц
Количество процессоров	2 (8 ядер/узел)	2 (8 ядер/узел)	2 (24 ядра/узел)
Пропускная способность канала доступа к памяти	12.8 ГБ/сек	32 ГБ/сек на проц.	42.7 ГБ/сек на проц.
Интерконнект	IB DDR	IB QDR	IB QDR

Внутриузловая масштабируемость, MPI



1 – один процессор
2 – два процессора

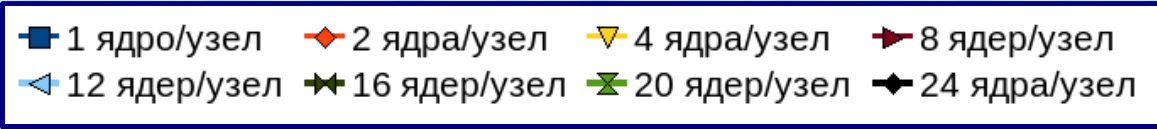
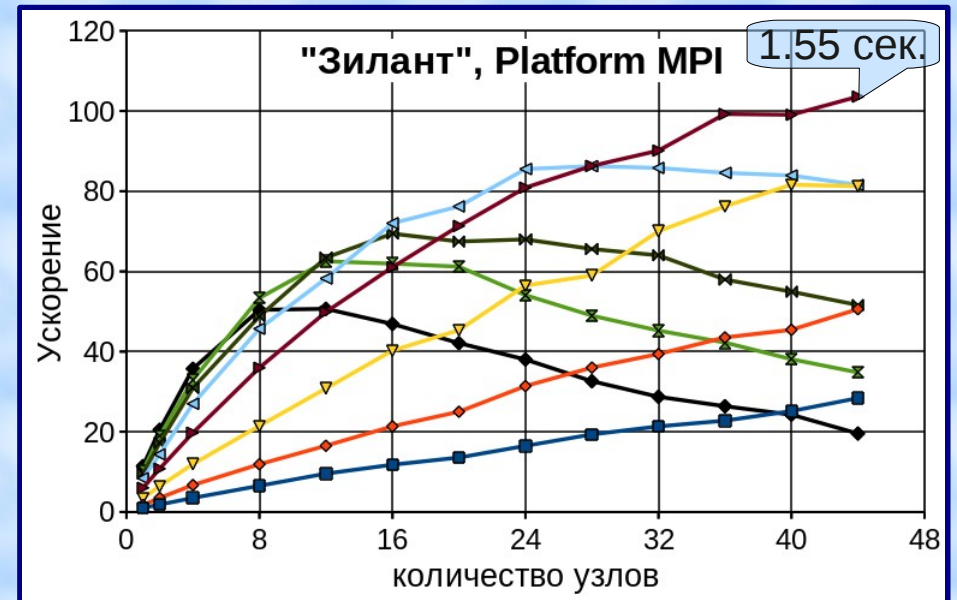
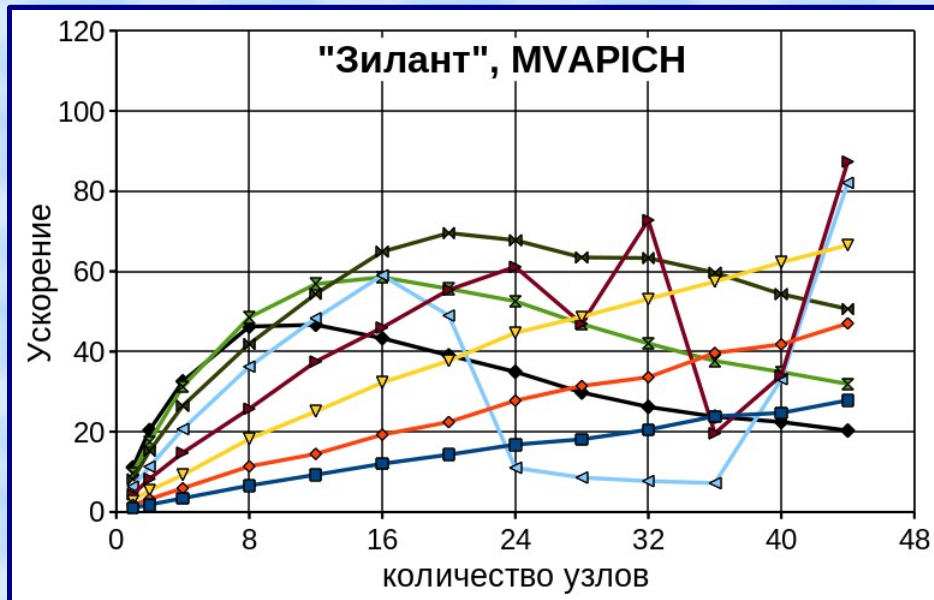
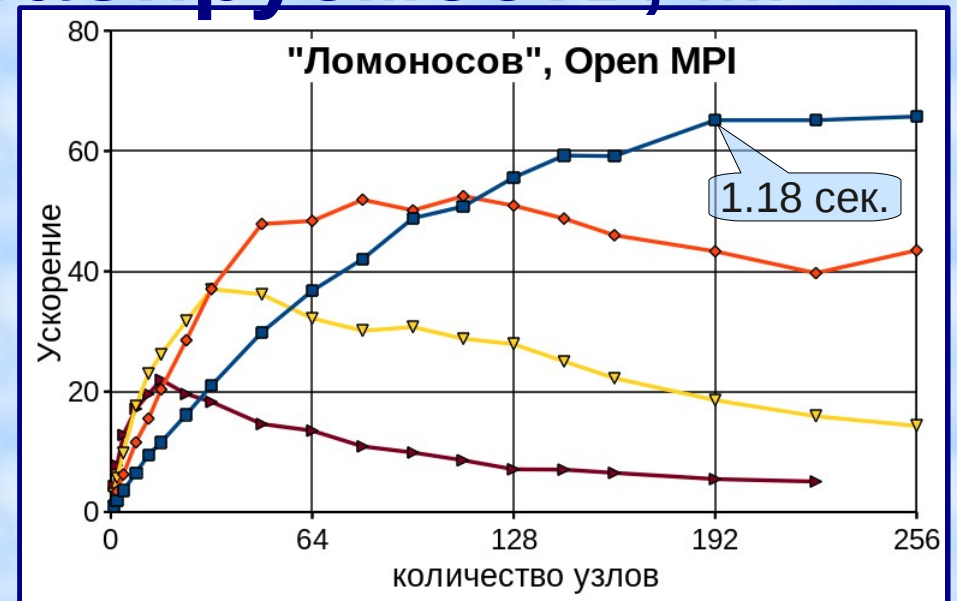
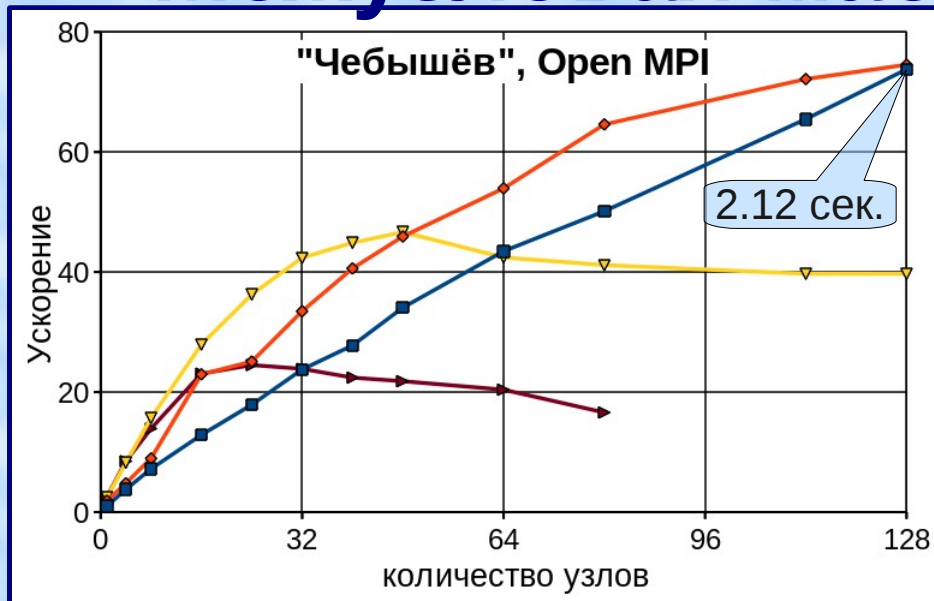
Внутриузловая масштабируемость, MPI+ShM

Вычислительная система	Размер блоков							
	1	2	3	4	6	8	12	24
“Чебышёв”	62	61.74	-	61.25	-	60.43	-	-
“Ломоносов”	16.13	17.04	-	22.6	-	21.1	-	-
“Зилант”	14.13	12.92	13.06	-	13.89	-	23.88	49.19

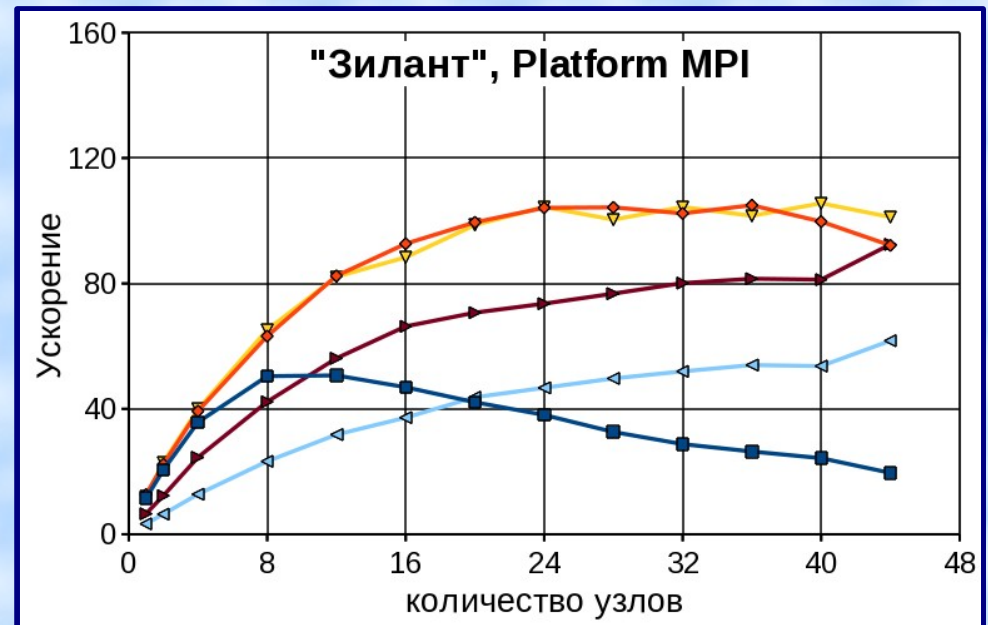
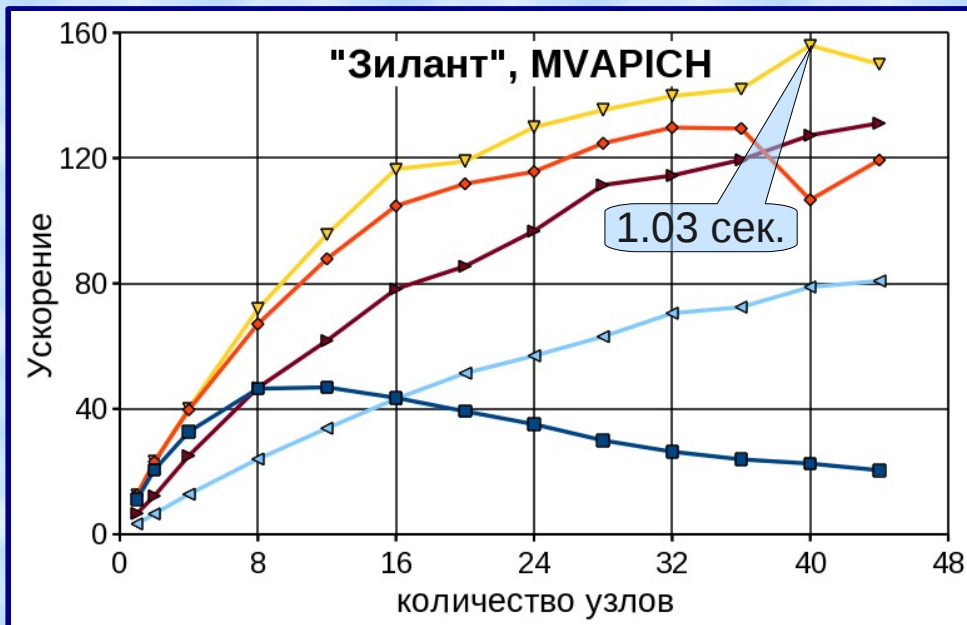
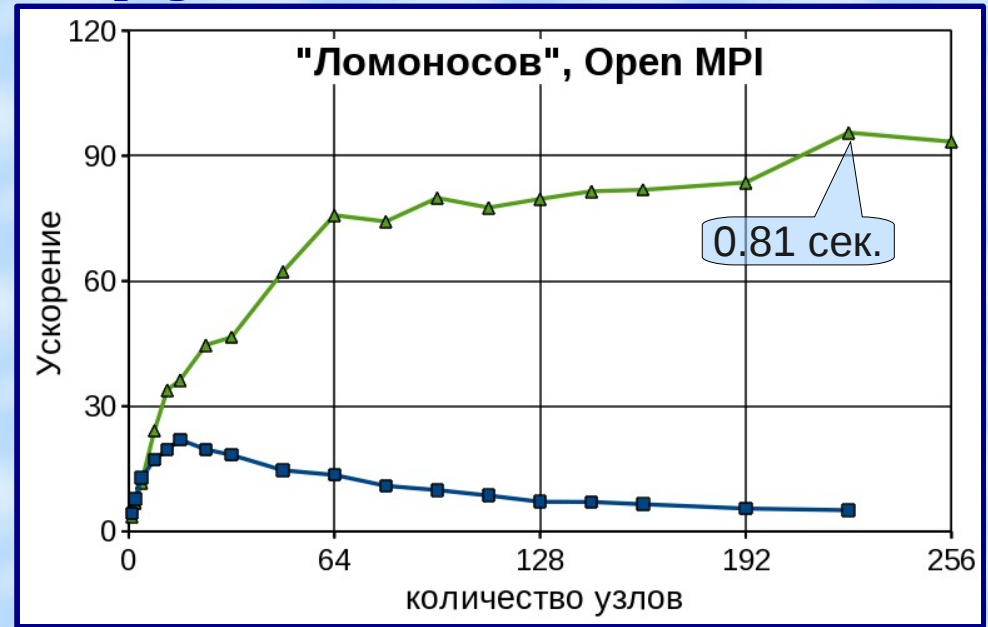
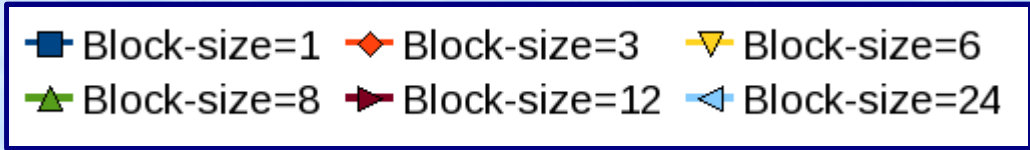
- Используются все имеющиеся ядра на узле
- Размер блока – количество агрегированных MPI-процессов
- Необходимо использовать “привязку” процессов к ядрам
- Оптимальная схема загрузки ядер на узле для AMD-платформы:

```
CPU_MAP:0,12,6,18,3,15,9,21,1,13,7,19,4,16,10,22,2,14,8,20,5,17,11,23
```

Межузловая масштабируемость, MPI



Межузловая масштабируемость, MPI+ShM



Доработанная версия гибридной модели

“Зилант”	Размер блоков					
	1	2	3	6	12	24
Исходный вариант	14.53	-	12.87	12.77	24.3	48.86
Модифицированный вариант	13.86	13.47	12.78	12.53	21.54	35.21

Заключение

- Проведенные тесты показали, что для эффективного использования многоядерных процессоров необходимо использовать привязку вычислительных процессов к физическим ядрам процессоров, причем при выборе схемы привязки следует исходить из архитектуры используемой аппаратной платформы.
- Выбор библиотеки MPI может существенным образом сказаться на масштабируемости приложения. Для MPI и для гибридной версий приложения наиболее удачный выбор библиотеки MPI может оказаться различным.
- При использовании только одного вычислительного узла MPI-приложения могут оказаться столь же эффективными, как и приложения, работающие через общую память.
- При реализации гибридных моделей на NUMA-архитектуре необходимо минимизировать размер общих для всех процессов областей памяти, чтобы сократить количество обращений в память другого NUMA-node через межпроцессорный интерфейс.

Благодарности

Приведенные в работе результаты тестирования были получены на:

- вычислительных системах **СКИФ МГУ «Чебышёв»** и **«Ломоносов»** суперкомпьютерного комплекса МГУ им. М.В. Ломоносова
- вычислительной системе **«Зилант»** ЗАО «Т-Сервисы»

