

**НОЦ “Суперкомпьютерные технологии” МГУ**

**Семинар**

**"Суперкомпьютерные технологии в науке, образовании и промышленности"**

**1 апреля 2014 г.**

**Расчеты лазерной абляции и другие  
крупномасштабные атомистические модели:  
максимизация параллельной эффективности на  
кластере "Ломоносов"**

**В.В. Писарев, Г.С. Смирнов, С.В. Стариков, В.В. Стегайлов**

**Объединенный институт высоких температур РАН**

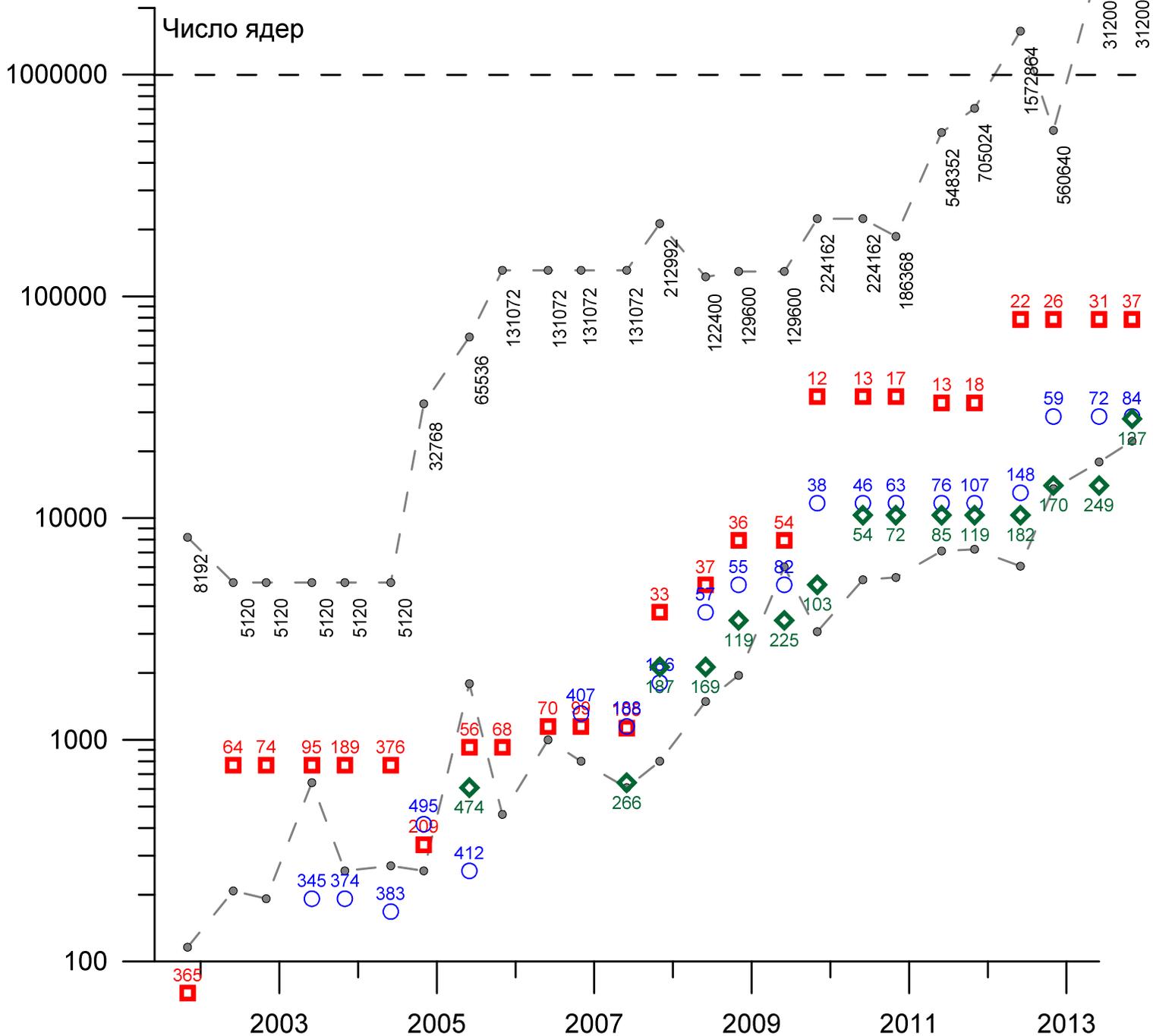


# СОДЕРЖАНИЕ

1. Движение в «эру эксафлопса»
2. Параллельные методы в атомистических расчетах
3. Моделирование лазерной абляции
4. Результаты наших тестов производительности на «Ломоносове»

Движение в «эру эксафлопса»

# Суперкомпьютеры



# «MPP» vs «SMP»

Особенности устройства суперкомпьютеров №№1-3  
текущего списка Top500 2013 г.

	Кол-во узлов	Кол-во вычисл. ядер ЦПУ	Узел		
			ЦПУ (кол-во ядер)	Ускоритель (кол-во ядер)	Суммарная теор. произв. узла, ТФлопс
Tianhe-2	16 000	384 000	2 x Intel Xeon (2 x 12)	3 x Intel Phi (3 x 57)	3.43
Cray Titan	18 688	299 008	AMD Opteron (16)	Nvidia Kepler (14 потоковых мультипроцессора)	1.45
IBM Sequoia	98 304	1 572 864	IBM PowerPC (16 + 2)		0.205

Нужен ли в действительности  
массовый паралелизм?

# US DOE INCITE initiative

Год	Всего проектов	Проекты по молекулярному и атомистическому моделированию	% от общего числа
2006	15	6	40%
2007	45	11	24%
2008	58	16	28%
2009	66	20	30%
2010	70	23	33%
2011	57	18	32%
2012	60	23	38%
2013	61	24	39%

## 2012 год:

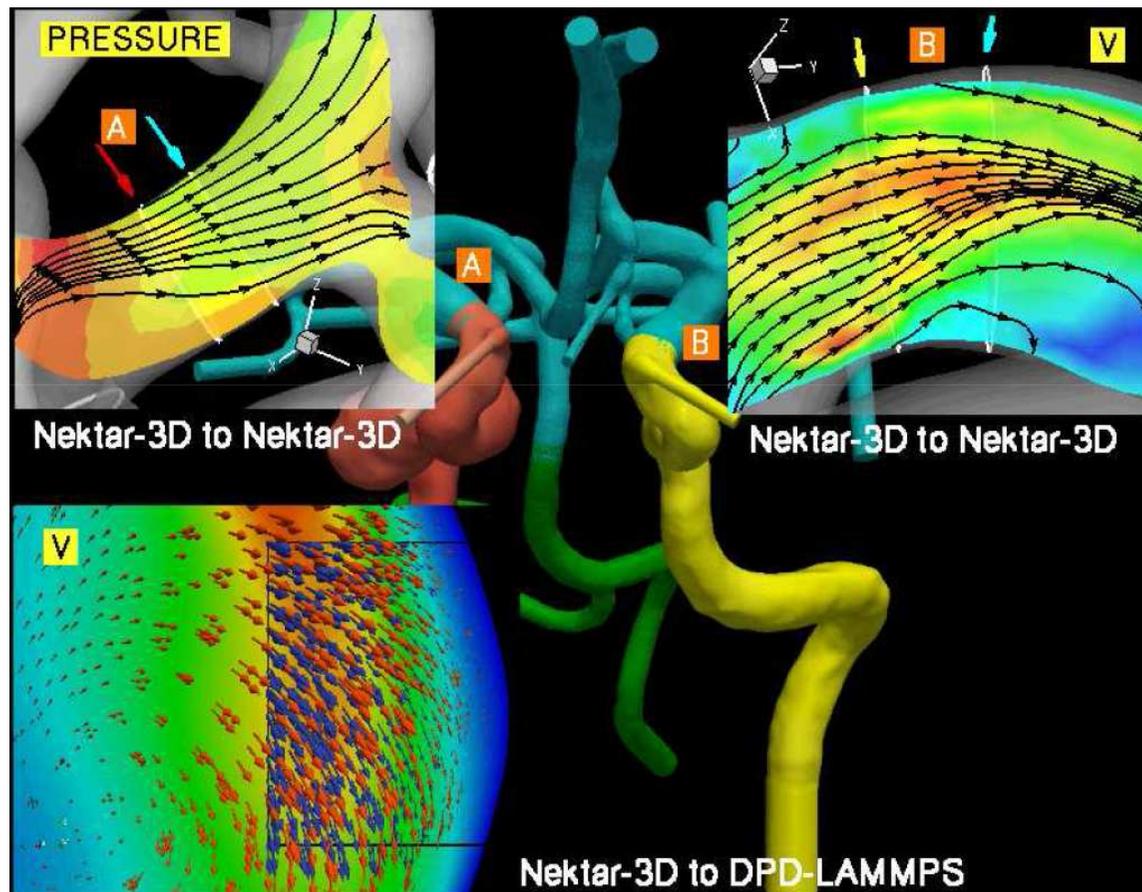
60 проектов = 1672 миллиона процессор-часов на IBM BlueGene/P и Cray XT5

23 проекта по молекулярному и атомистическому моделированию = 518 млн. час.

31 % общего времени

# US DOE INCITE initiative

Пример результатов комбинированного расчета кровотока с эффектами отложения холестериновых бляшек в рамках многомасштабной модели, сочетающей атомистическое моделирование с моделью жидкости Навье-Стокса



*Grinberg L., et al. A new computational paradigm in multiscale simulations: Application to brain blood flow // Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC), 2011.*

# US DOE INCITE initiative

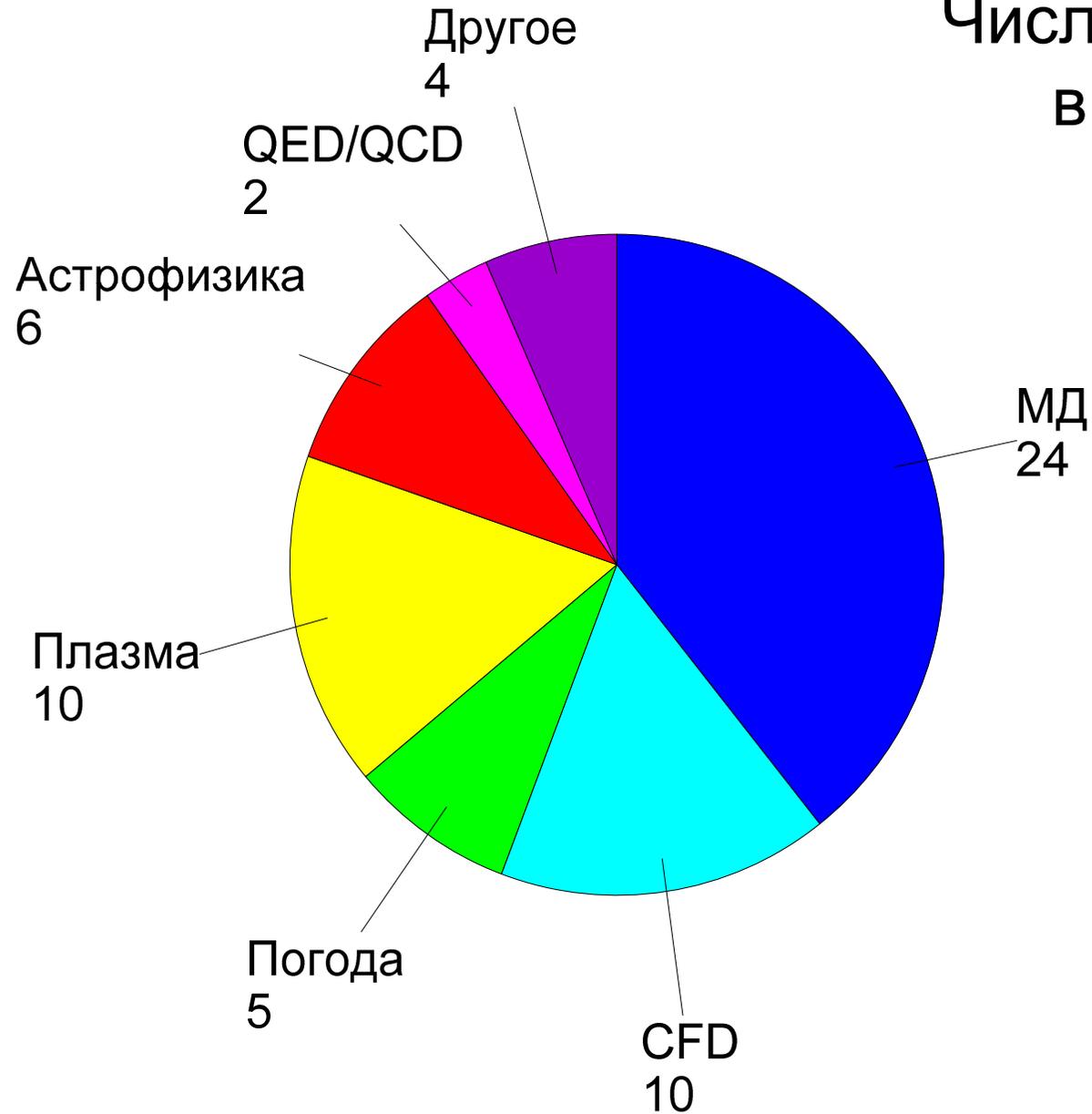
Демонстрация сохранения эффективности распараллеливания решения одной задачи на системах BlueGene/P и Cray XT5

Число ядер	Время загрузки ЦПУ, с	Эффективность
BlueGene/P (4 ядра на узел)		
28 672	3205.58	
61 440	1399.12	1.07
126 976	665.79	1.02
Cray XT5 (12 ядер на узел)		
17 280 (Kraken)	2194	
25 920 (Kraken)	1177	1.24
34 560 (Jaguar)	806	1.10
93 312 (Jaguar)	280	1.07
186 624 (Jaguar)	206	0.68

*Grinberg L., et al.* **A new computational paradigm in multiscale simulations: Application to brain blood flow** // Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC), 2011.

# US DOE INCITE initiative

Число проектов  
в 2013 г.



# Пример из программы SC'13

**TITLE: Petascale Direct Numerical Simulation of Turbulent Channel Flow on up to 786K Cores**

SESSION: Extreme-Scale Applications

EVENT TYPE: Papers, Awards, Best Student Paper Finalists

TIME: 1:30PM - 2:00PM

SESSION CHAIR: Michael A. Heroux

AUTHOR(S): Myoungkyu Lee, Nicholas Malaya, Robert D. Moser

ROOM:201/203

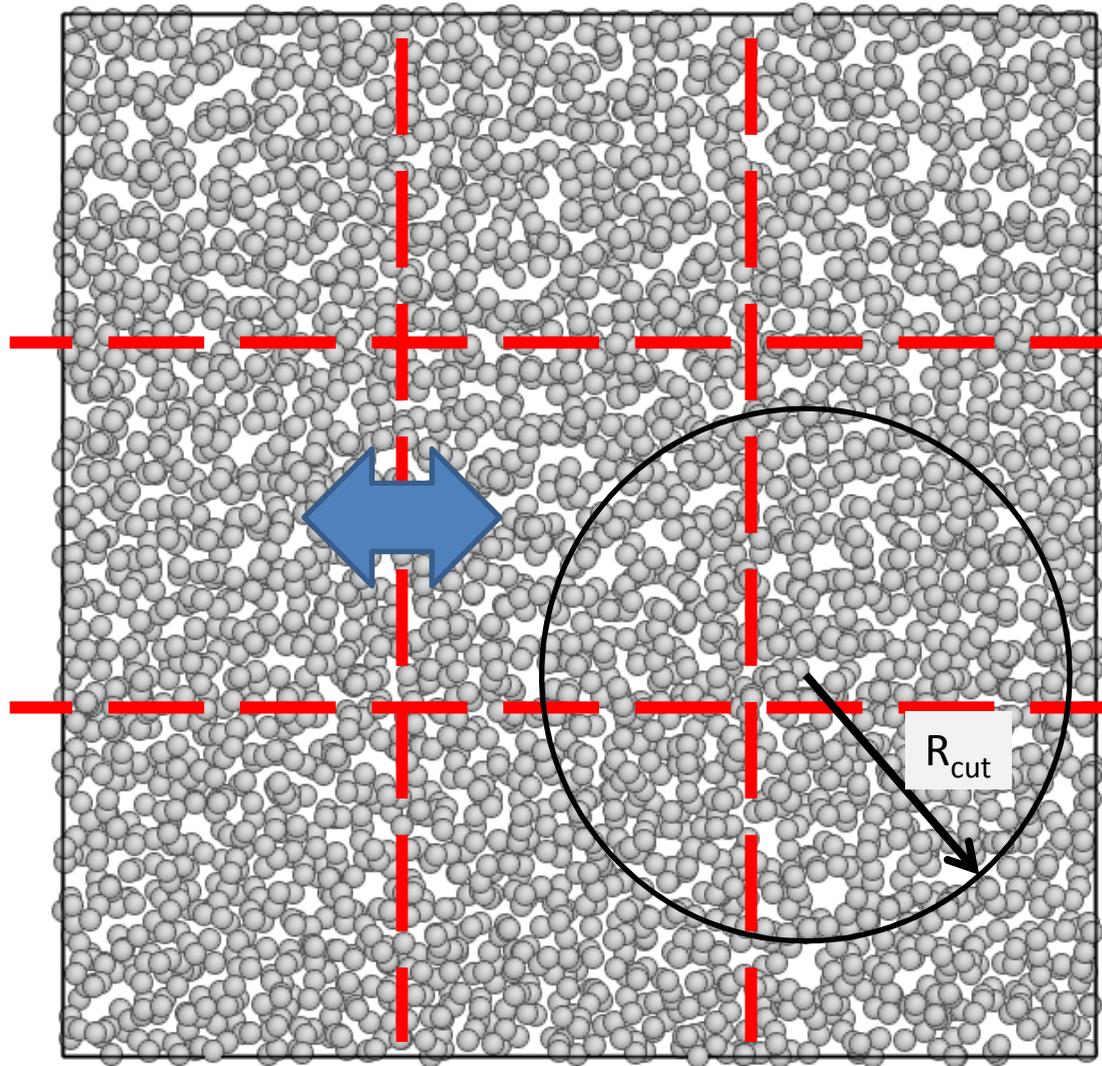
ABSTRACT:

We present results of performance optimization for direct numerical simulation (DNS) of **wall bounded turbulent flow (channel flow)**. DNS is a technique in which the fluid flow equations are solved without subgrid modeling. **Of particular interest are high Reynolds number (Re) turbulent flows over walls**, because of their importance in technological applications. Simulating high Re turbulence is a challenging computational problem, due to the high spatial and temporal resolution requirements. An optimized code was developed using **spectral methods**, the method of choice for turbulent flows. Optimization was performed to address three major issues: efficiency of banded matrix linear algebra, cache reuse and memory access, and communication for the global data transposes. Results show that performance is highly dependent on characteristics of the communication network, rather than single-core performance. In our tests, it exhibits approximately **80% strong scaling parallel efficiency at 786K cores relative to performance on 65K cores**.

Эти цифры демонстрируют  
переход к новой эре  
использования  
вычислительных методов  
в науке и технике

# Параллельные методы в атомистических расчетах

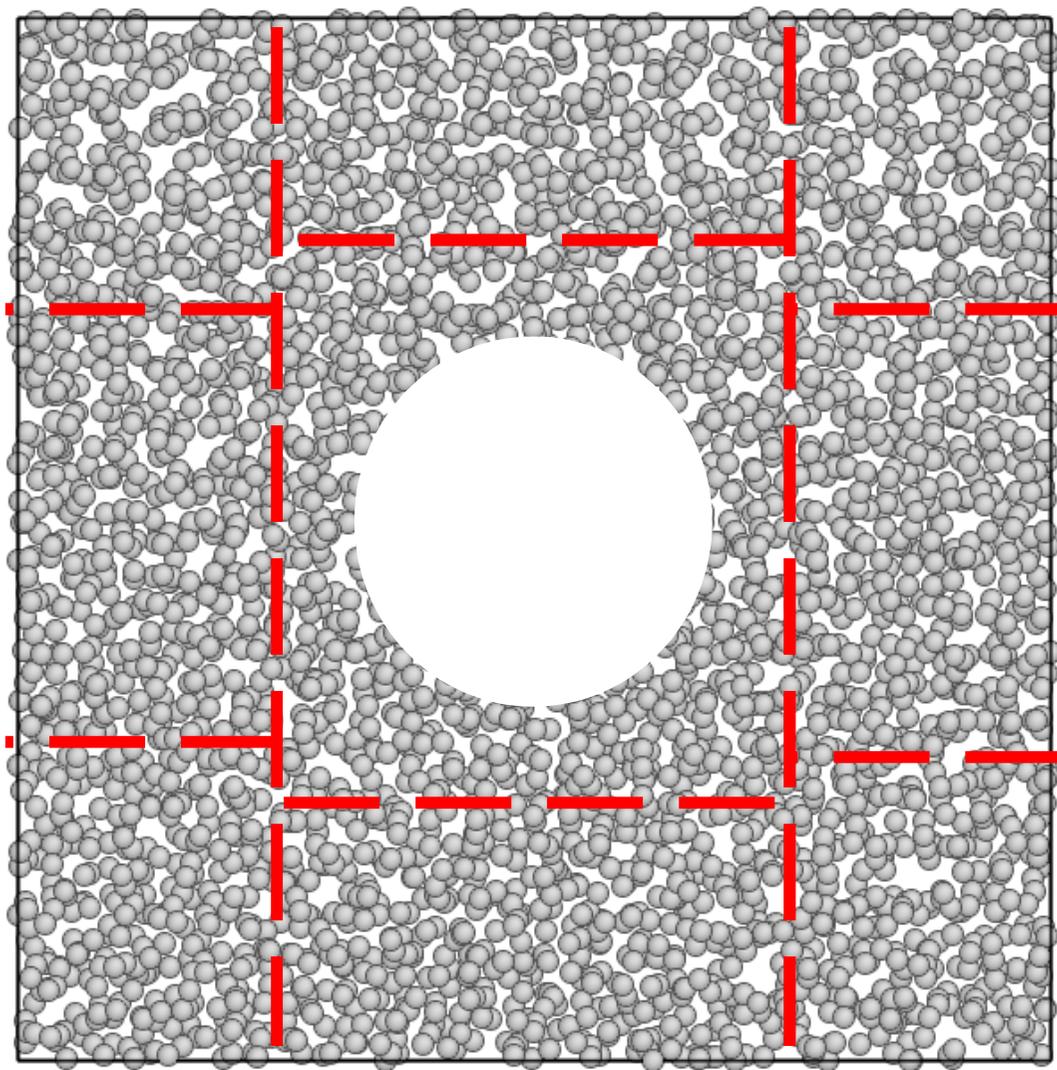
# Декомпозиция по пространству



$$m_i \frac{d^2 r_i}{dt^2} = - \sum_j \frac{\partial U_{ij}}{\partial r_{ij}} \rightarrow \begin{matrix} r_i(t) \\ v_i(t) \end{matrix}$$

Рекорд по числу атомов  $4,125 \times 10^{12}$

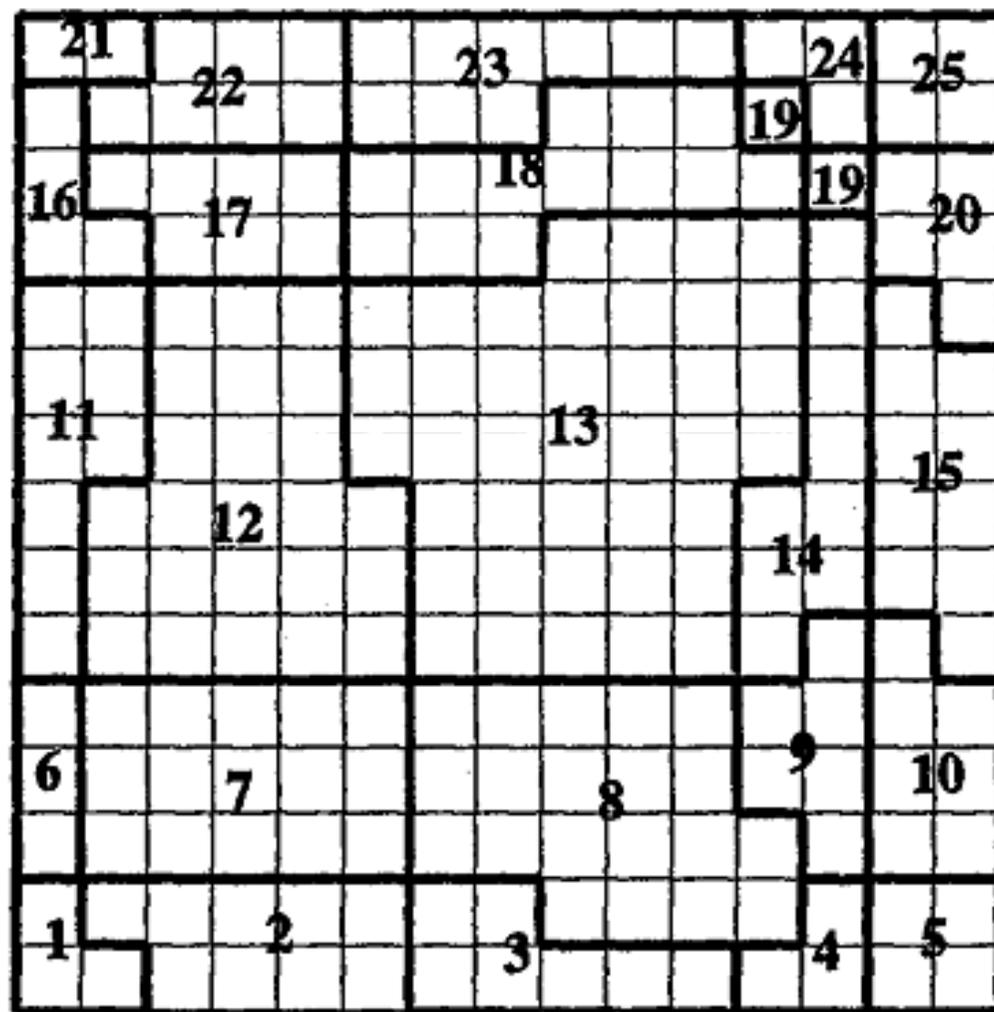
## Динамическая балансировка нагрузки



Минимизация соотношения

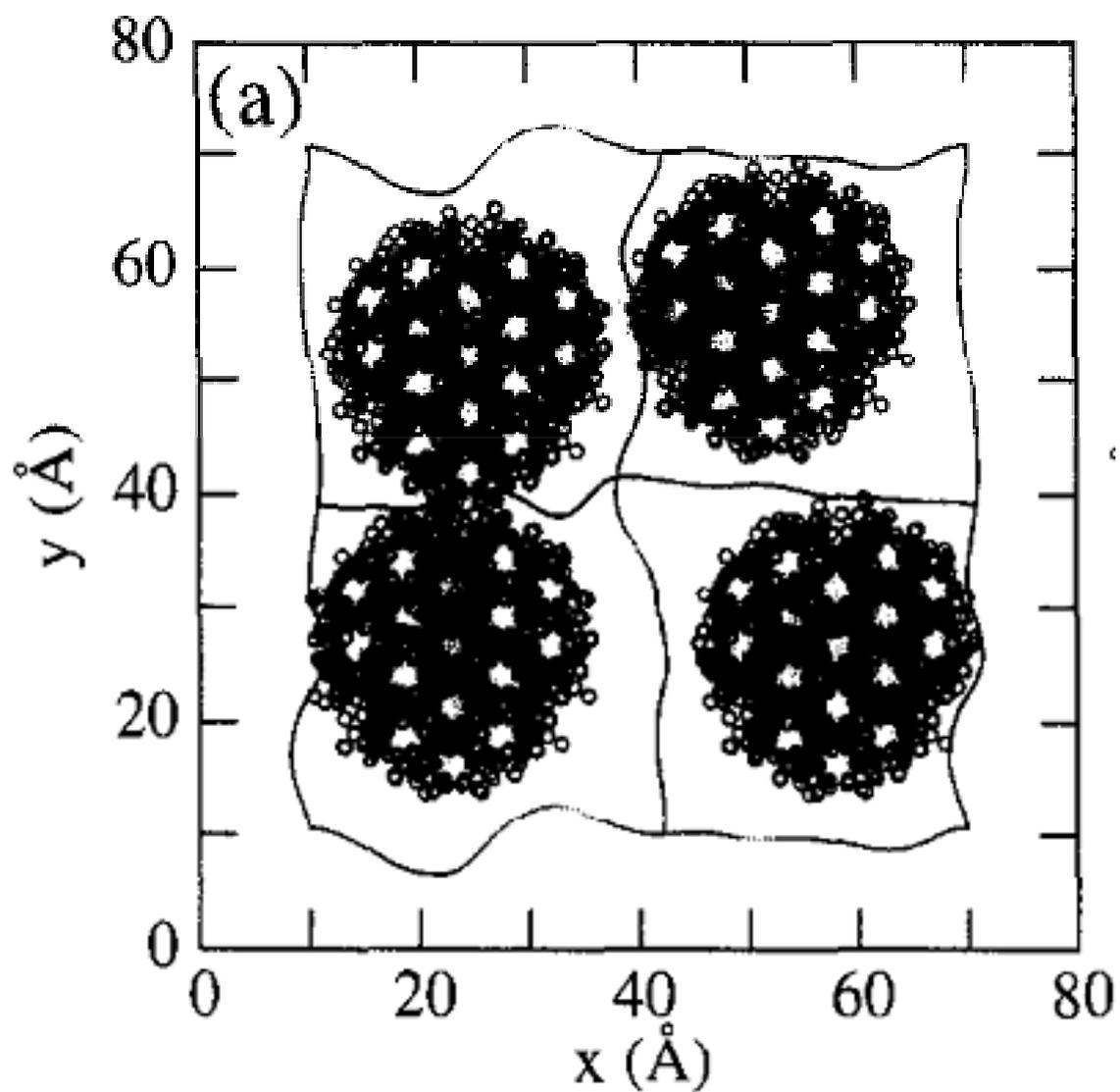
$$N_{\max}/N_{\text{ave}}$$

## Динамическая балансировка нагрузки

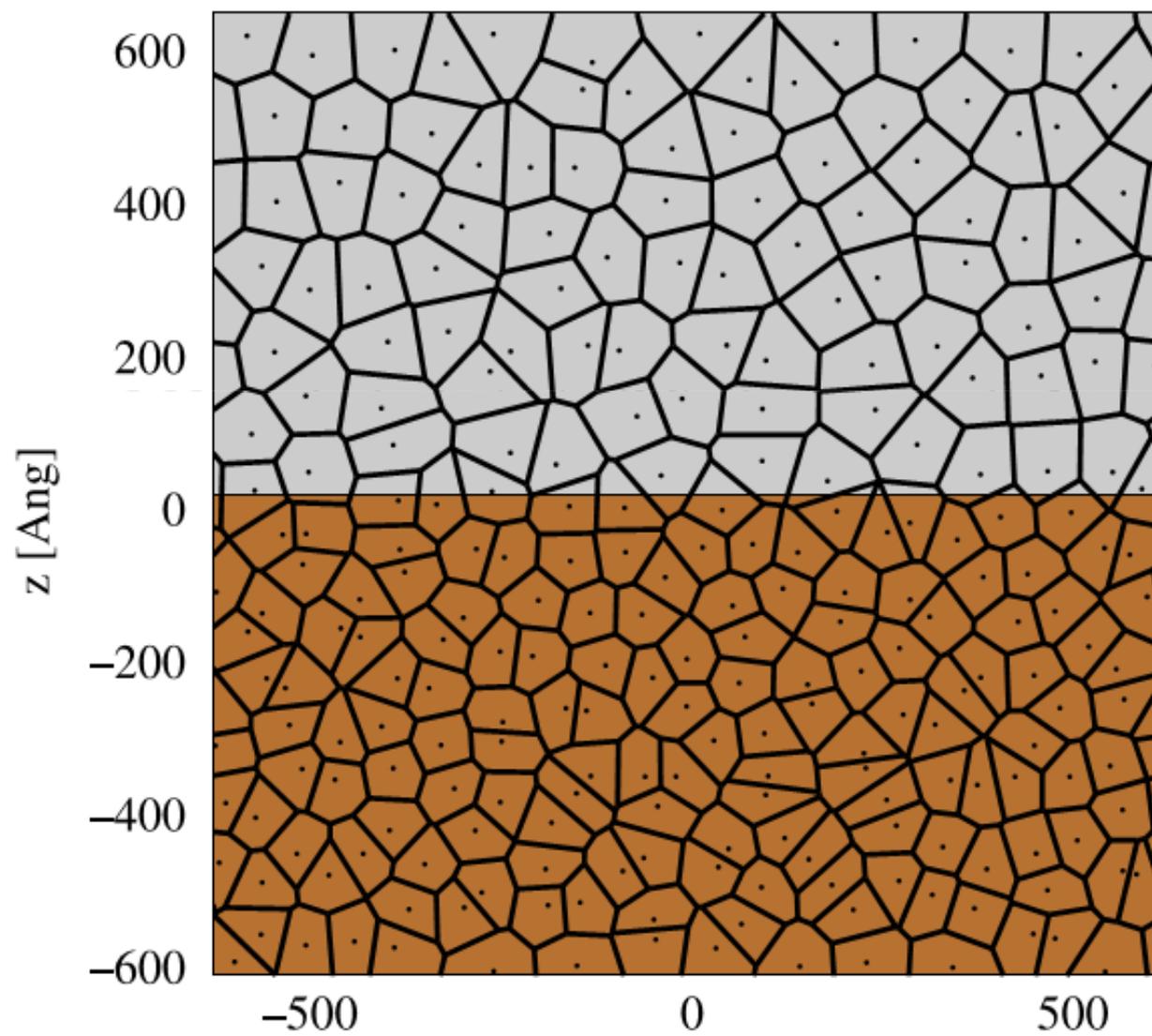


△ (13)

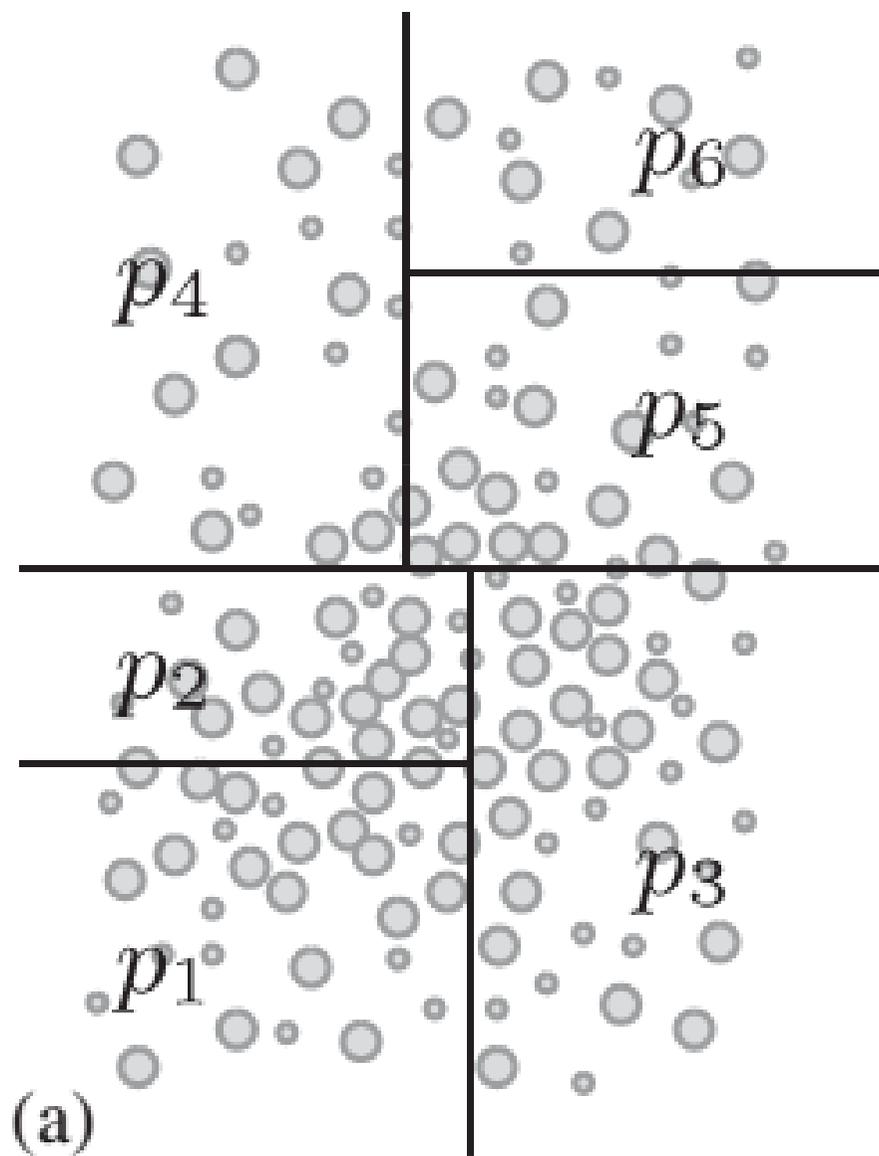
## Динамическая балансировка нагрузки



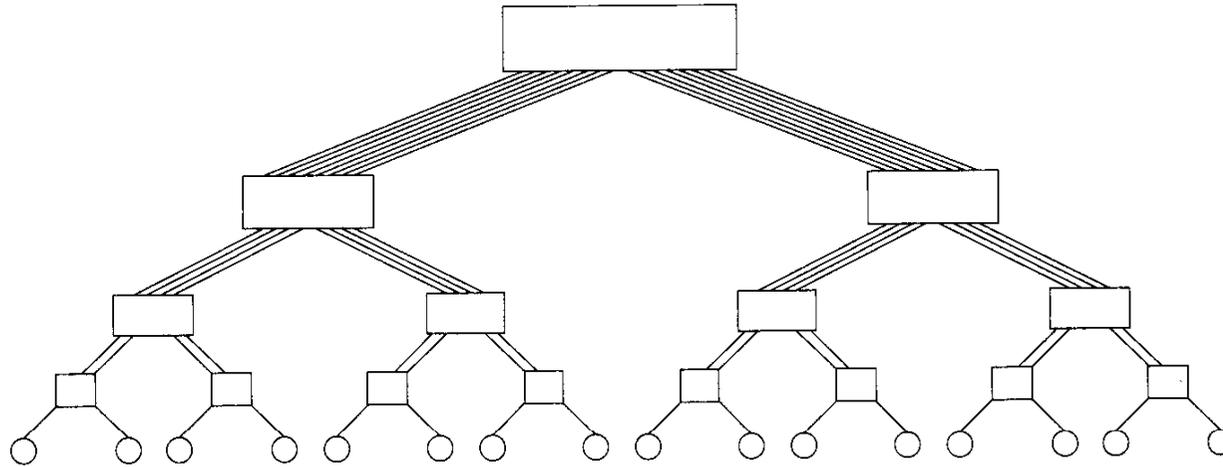
## Динамическая балансировка нагрузки



## Динамическая балансировка нагрузки



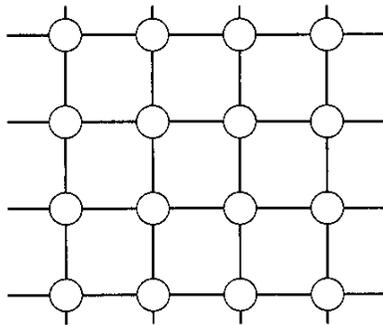
# Топология интерконнекта и архитектура



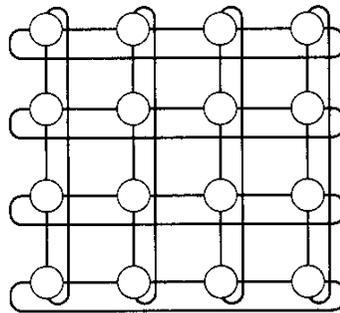
**OpenMP**  
**CUDA**  
**OpenCL**  
**ПЛИС**

....

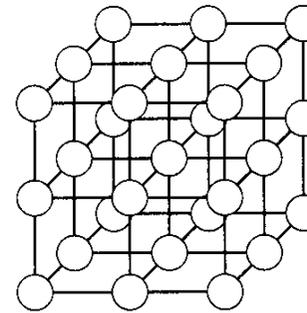
## Message Passing Interface



(a)



(b)



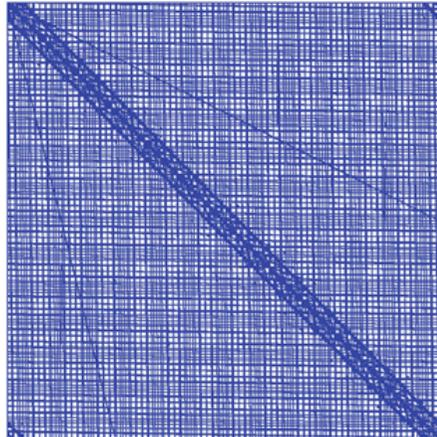
(c)

# Тороидальные системы в мире

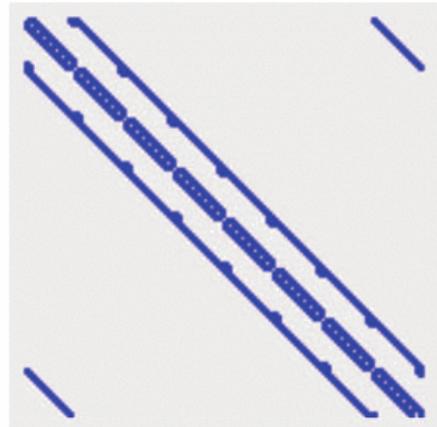
Регион/страна	Кол-во	Число ядер	$\Sigma R_{\max}$ , ТФлопс	% от $\Sigma R_{\max}$	Название
Россия	1	8192	23.9	0.01	IBM – 1.
Европа	17	1 406 240 (+45 712)	14 371.6	6.4	IBM – 8, Cray – 9.
Австралия, Бразилия, Канада, Сауд.Аравия	4	194 944 (- 8 192)	1 465.8	0.65	IBM – 3, Cray – 1.
США	35	4 810 064 (+367 416)	59 812.9	26.7	IBM – 18, Cray – 16, Sun – 1.
Япония и Южная Корея	10	1 042 688 (+23 808)	14 202.9	6.3	Fujitsu -3, IBM – 2, Cray – 5.
Итого:	67	7 462 128 (+428 744)	89 877.1	40.1	Fujitsu -3, IBM – 32, Cray – 31, Sun – 1.

Число и суммарные параметры систем с тороидальной топологией по странам мира в списке Top500 июня 2013 г.: количество подобных систем, суммарное число ядер (в скобках – прирост по отношению к предыдущему списку), суммарная производительность, вклад в суммарную производительность списка июня 2013 г. (224 ПФлопс), распределение систем по производителям. (Китай пока не представлен в Top500 такими суперкомпьютерами)

# Модель коммуникации различных приложений

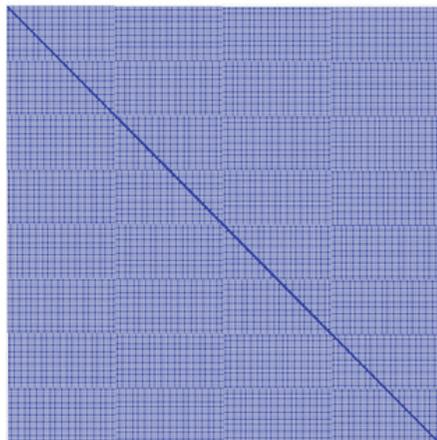


(a) NAMD\_PME 256K

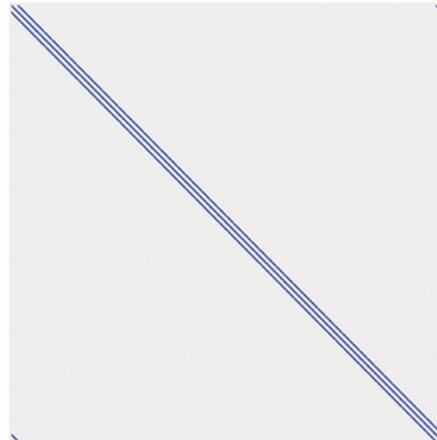


(b) MILC 4K

(a) NAMD  
(b) MILC



(c) CG 64K

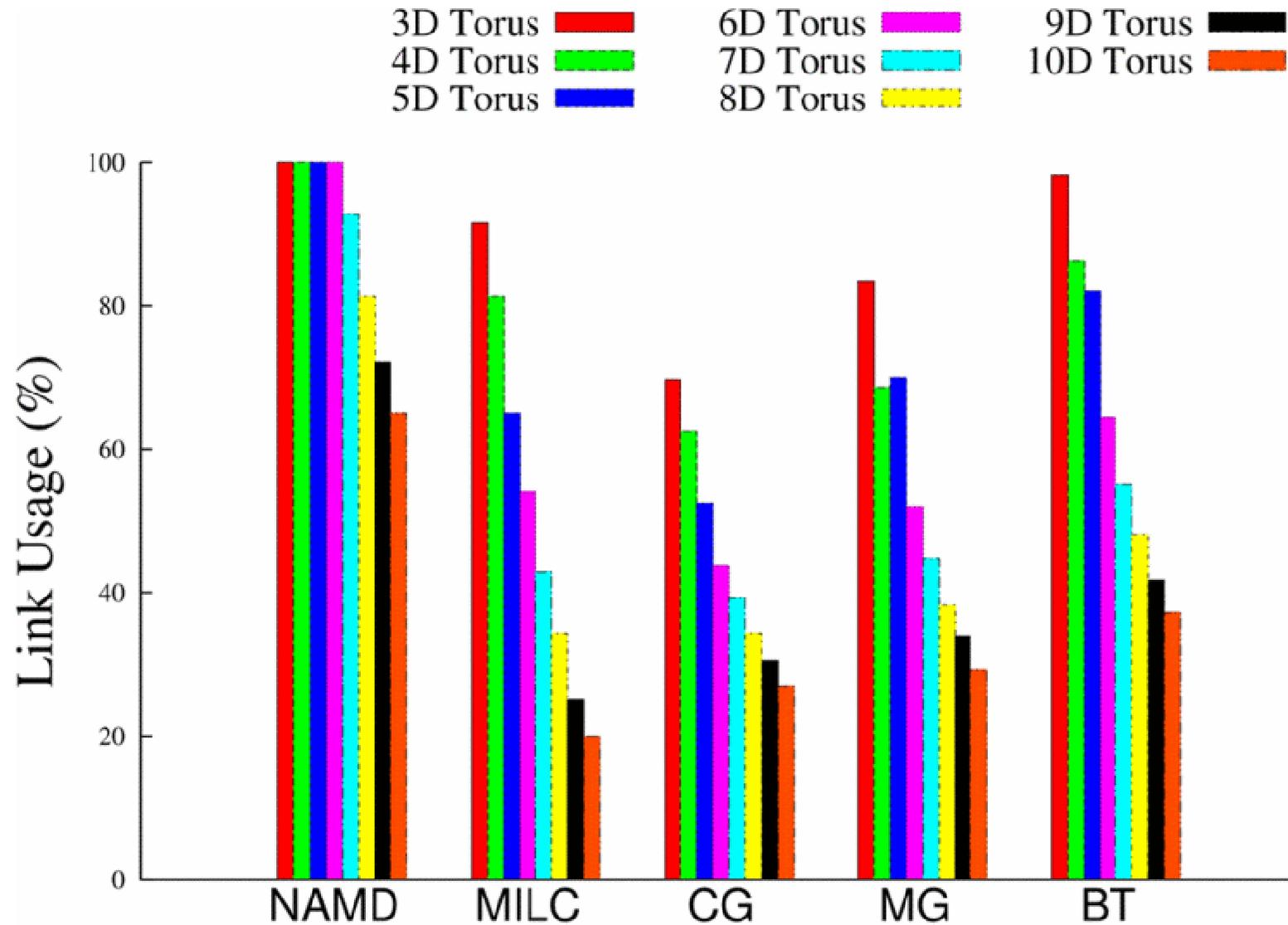


(d) MG 64K

NAS Parallel Benchmarks:  
(c) CG – Conjugate Gradient  
(d) MG – MultiGrid

E.Totoni, N. Jain, L.V. Kale, "Toward Runtime Power Management of Exascale Networks by On/Off Control of Links", 2013

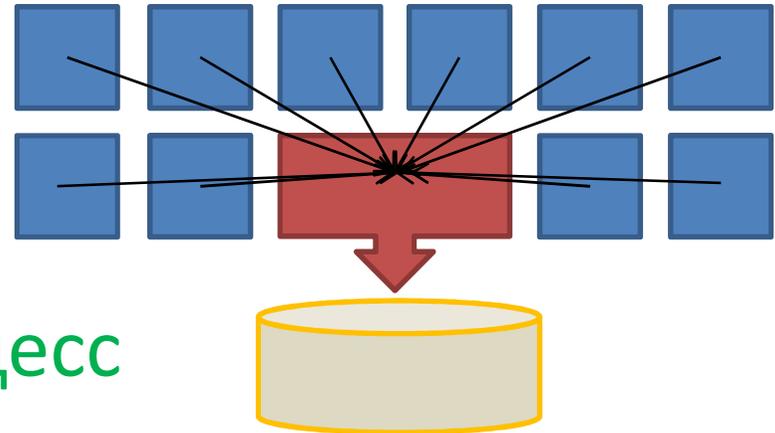
# Процент используемых связей интерконнекта



## Ввод-вывод больших файлов

- **Сохранение** данных для анализа
- **Загрузка** начальных условий
- **Создание** контрольных точек

## Последовательный вывод



Вывод выполняет один процесс

### Проблемы:

- Схема не масштабируется
- Время растёт с увеличением
  - числа процессоров
  - количества данных

## Нужен ли параллельный ввод-вывод?

Для системы из  $10^9$  атомов минимальный размер файла  $7 \times 8 \text{ bytes} \times 10^9 = 52 \text{ Гб}$

id x y z v<sub>x</sub> v<sub>y</sub> v<sub>z</sub>

### Последовательный ввод-вывод

- 16 ядер – 1% времени
- 1024 ядра – 39% времени

"A supercomputer is a device for converting a CPU-bound problem into an I/O bound problem."

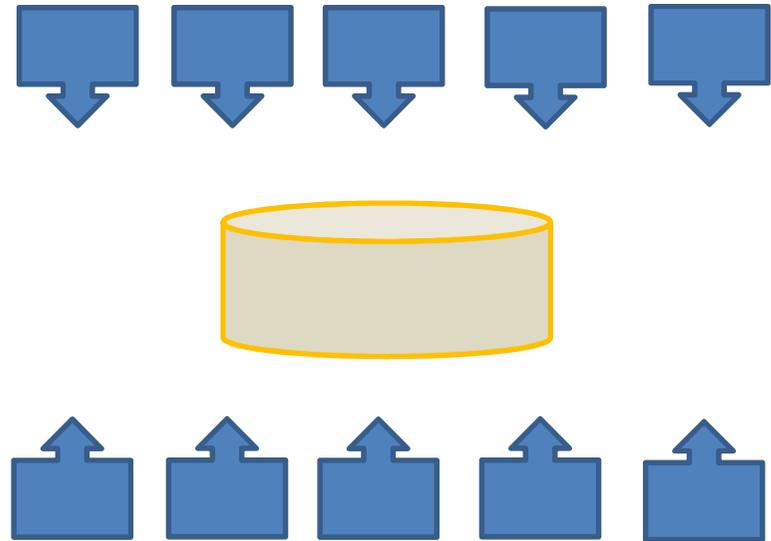
[Ken Batcher]

## Способы параллельного ввода-вывода

Один процесс – один файл

Ограничения:

- тип файловой системы
- нет масштабируемости
- число файлов



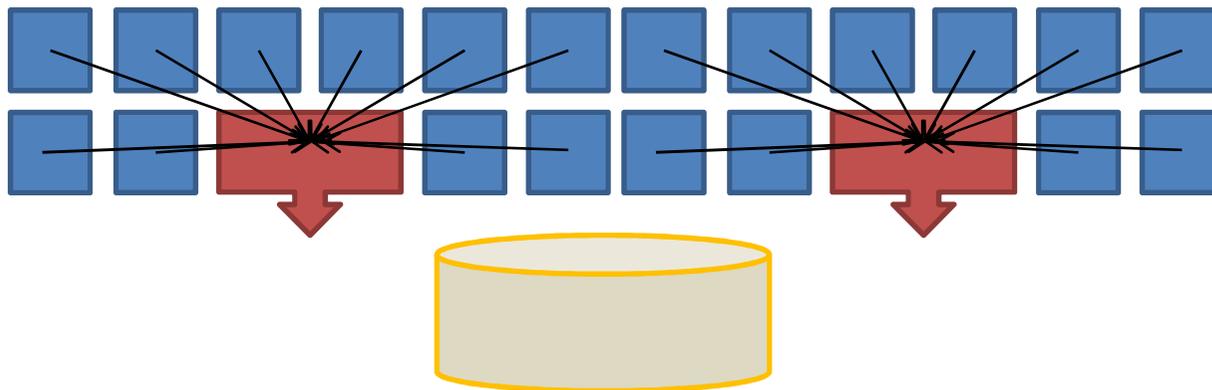
# Способы параллельного ввода-вывода

Вывод выполняют группы процессоров

Наиболее оптимальный способ!

Ограничения

перезапуск задач на другом числе процессоров





## Гибридные кластеры

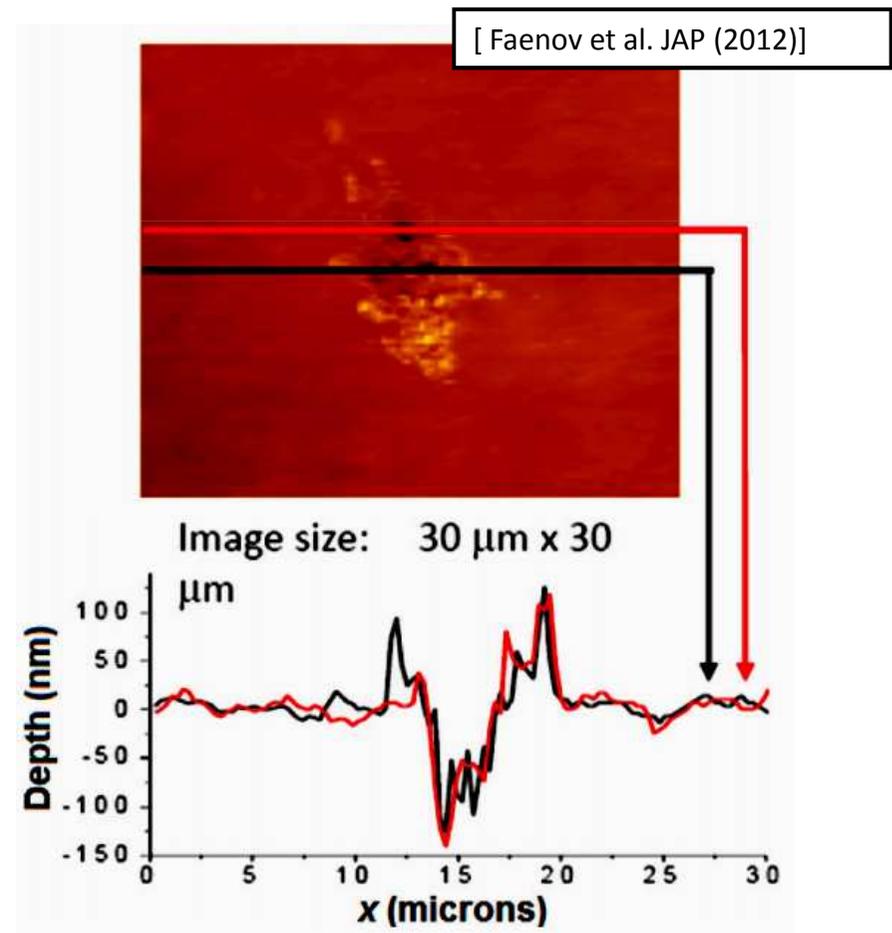
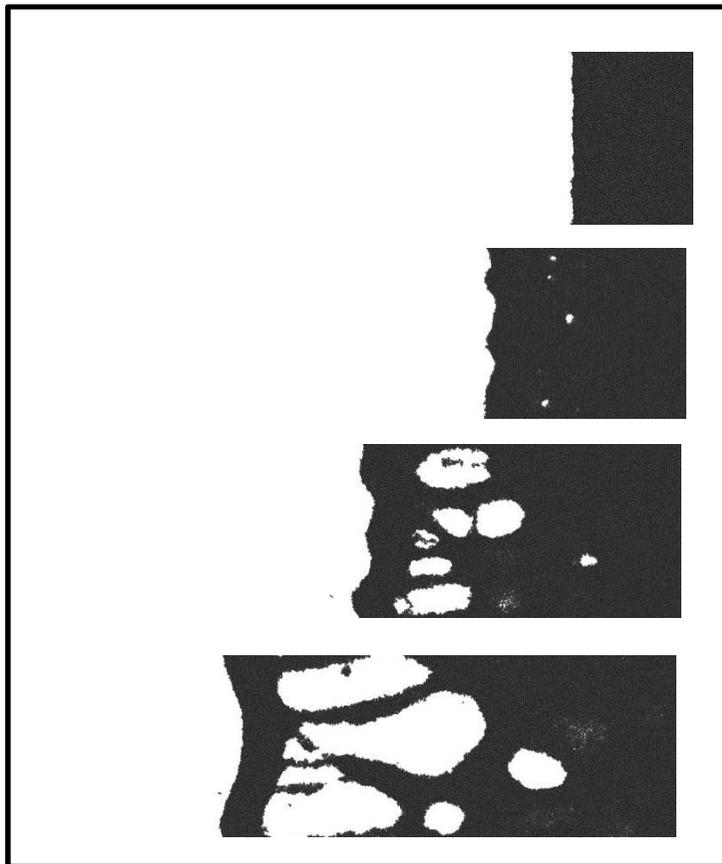
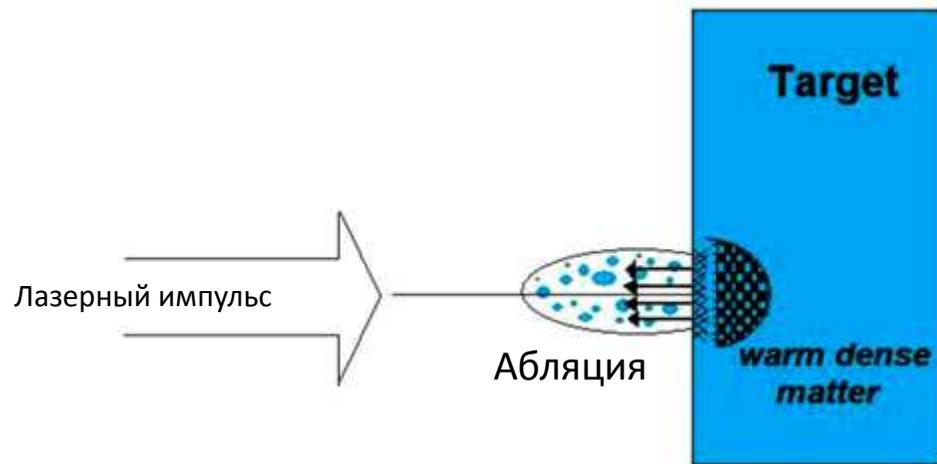


- Переход от скалярных вычислений к векторным
- Оптимизация инструкций gather-scatter
- Оптимизация загрузки гибридных элементов для совместного использования
- Ввод-вывод данных с ускорителей

## Выводы:

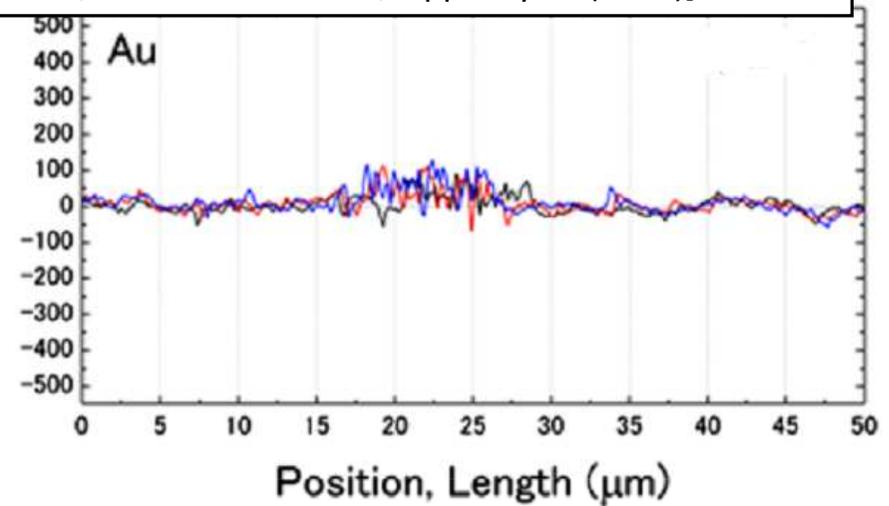
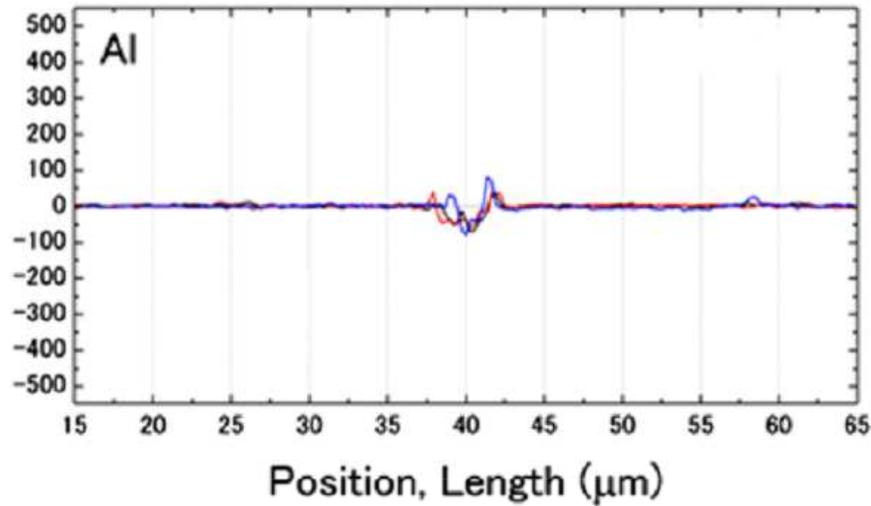
- Проанализированы проблемы распараллеливания МД задач на большом числе ядер
- Приведены аргументы, свидетельствующие о преимуществе тороидальных систем для данного класса задач

# Модификация поверхности лазерным импульсом

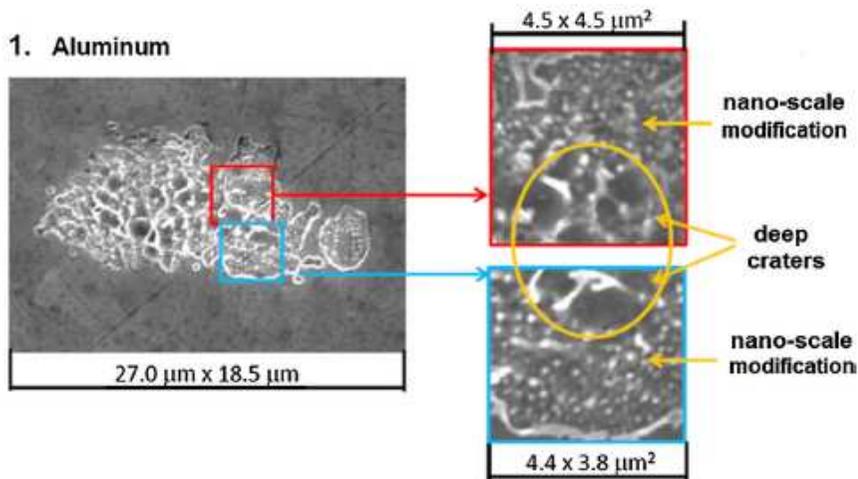


# Профиль поверхностной модификации при одинаковой энергии падающего импульса

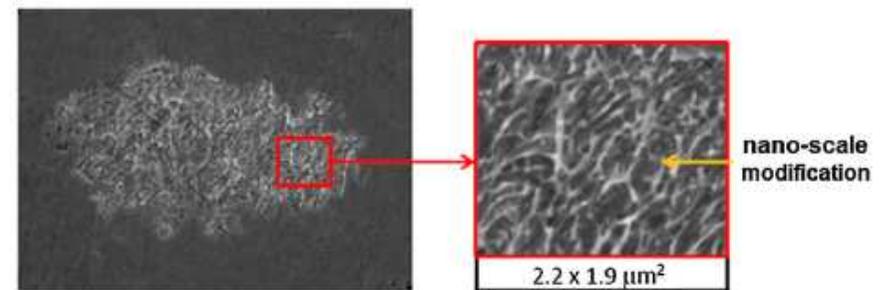
[M. Ishino, A.Yu. Faenov et al. / Appl Phys A (2013)]



## 1. Aluminum



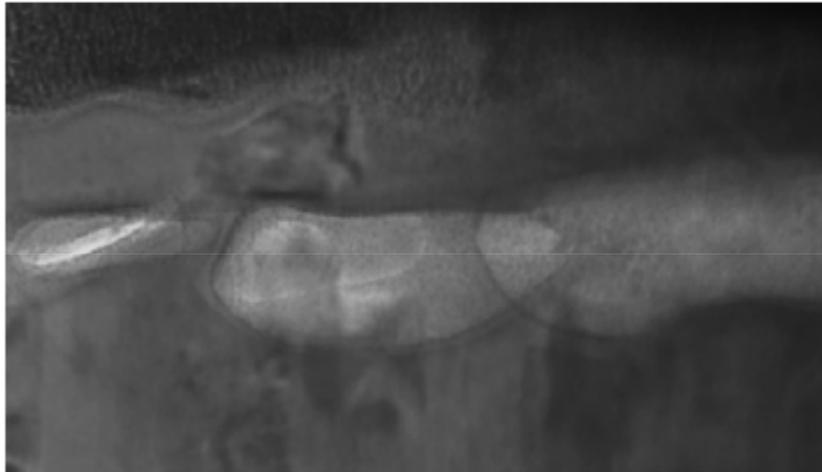
## 2. Gold



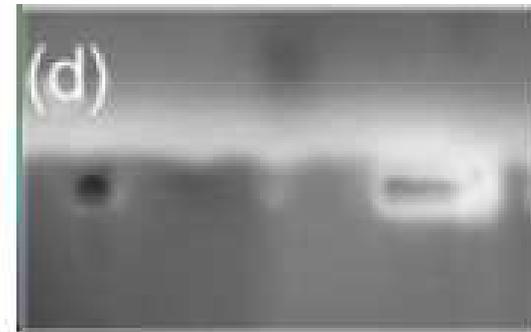
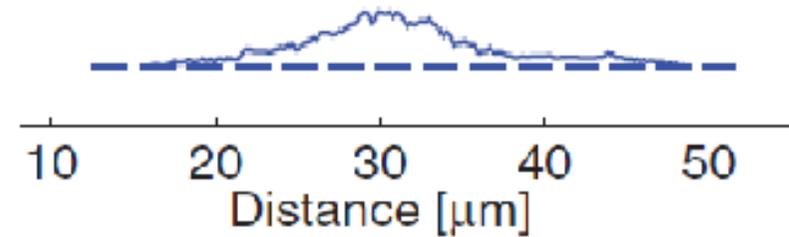
[S.V. Starikov, A.Yu. Faenov et al. / Appl Phys B (2014)]

# Распухание и образование нанополостей при энерговкладах $0.7 \cdot F_{abl} < F_0 / F_{abl} < 1$

Нанополости в алюминии



[Ашитков, Иногамов, Жаховский, Эмиров,  
Агранат, Олейник, Анисимов, Фортов //  
Письма в ЖЭТФ (2012)]



[Savolainen, Christensen,  
Balling // PRB (2011)]

# 2T-модель для моделирования абляции

ЭЛЕКТРОНЫ: уравнение теплопроводности (прибл. сплошной среды)  
 ИОНЫ: молекулярная динамика, электронное давление

Модификация открытого кода LAMMPS

$$C_e \frac{\partial T_e}{\partial t} = \nabla (K_e \nabla T_e) - g_p (T_e - T_i) + \nabla Q$$

$$m \frac{d\vec{v}_i}{dt} = \vec{F}_i (T_e) - \beta \vec{v}_i + \vec{\xi} (T_e) - \frac{\nabla P_e^{deloc}}{\rho_{ion}}$$

$g_p$  – factor of electron-ion relaxation

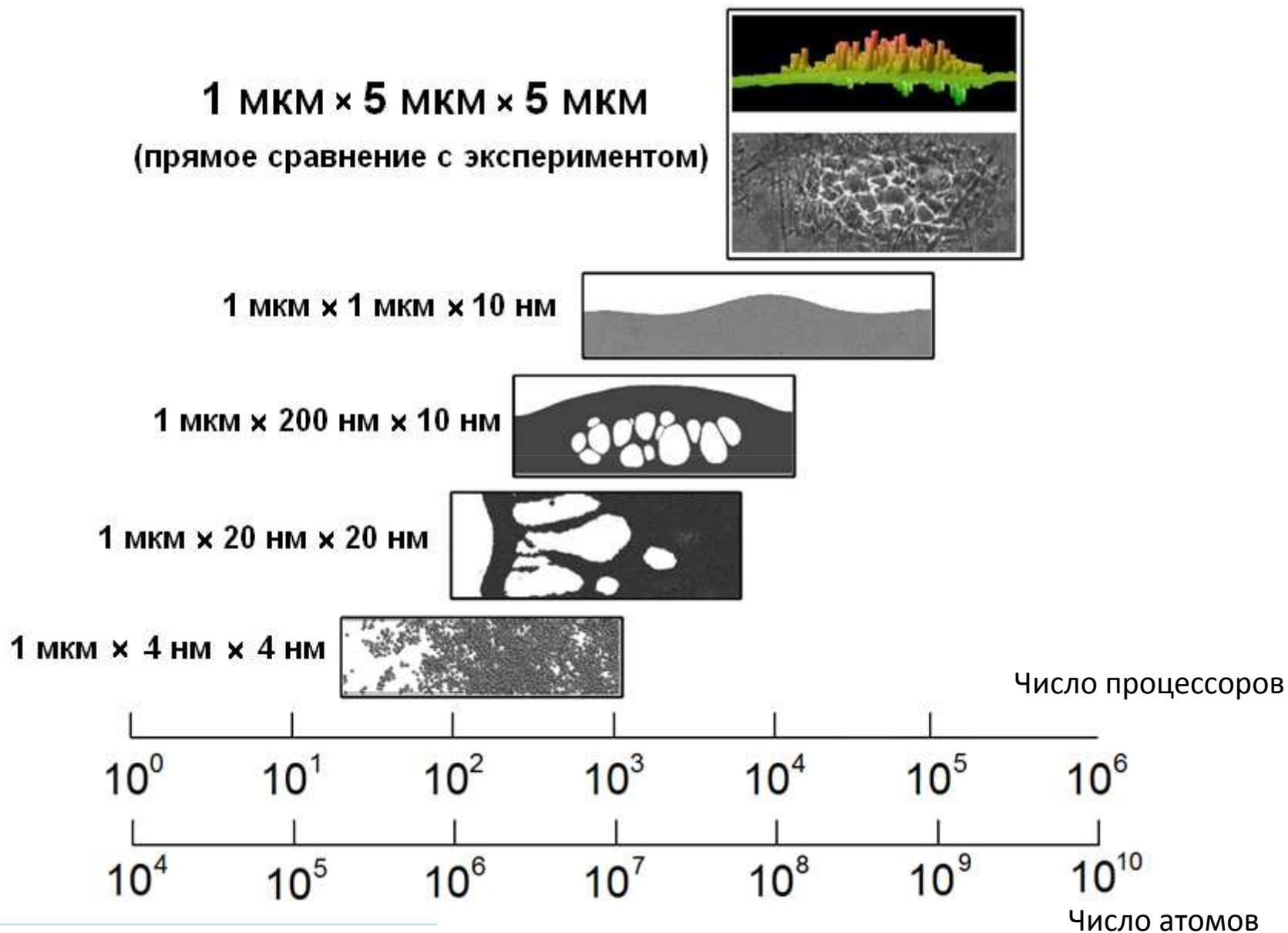
$C_e$  – electron heat capacity

$K_e$  – electron heat conductivity

$$\beta = \beta(g_p, C_e) \quad \xi \sim T_e^{1/2}$$

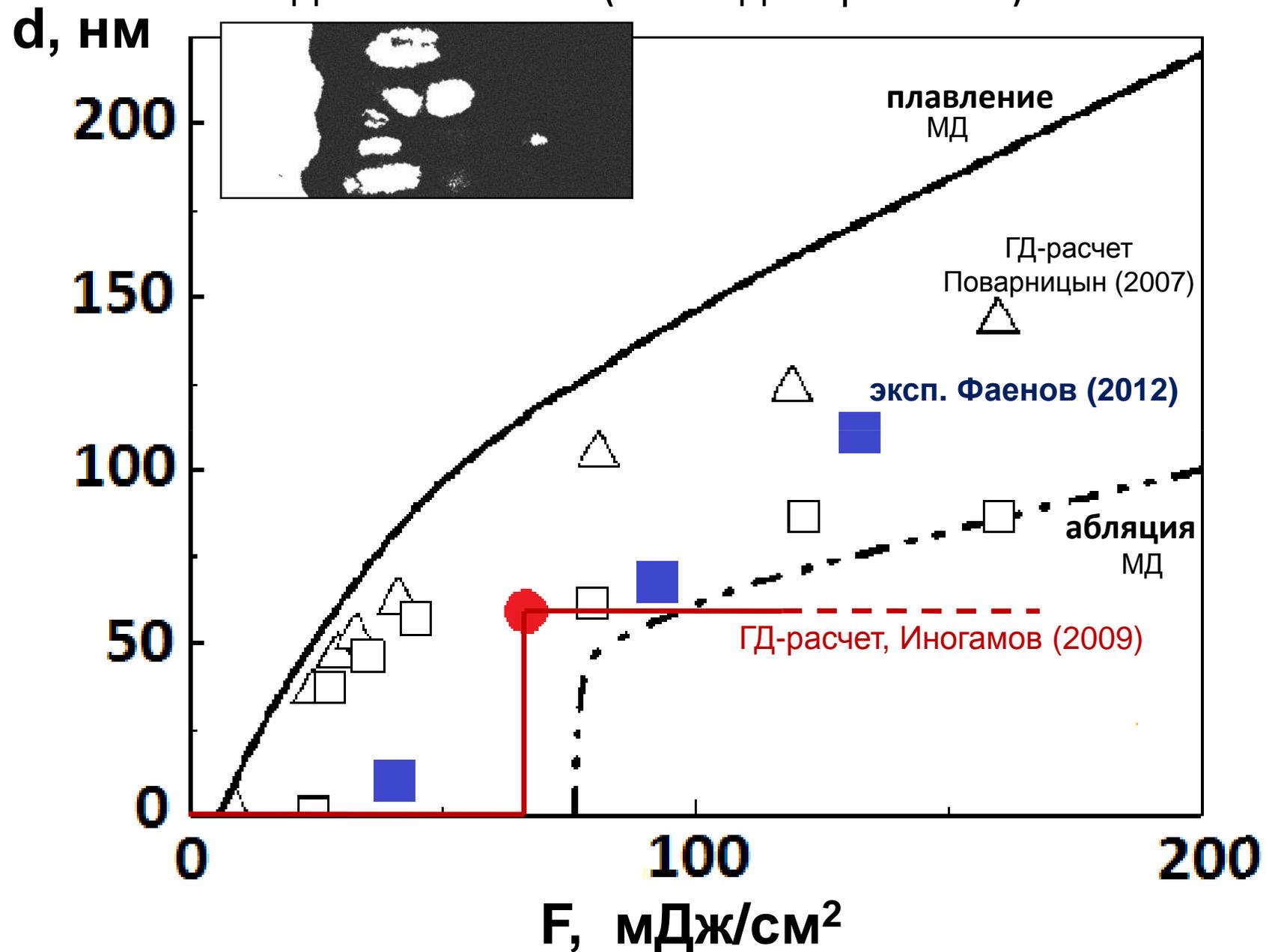
$$\tau = \frac{m}{\beta} \quad \tau \sim 10 \text{ ps}$$

# Крупномасштабное атомистическое моделирование наноструктурирования поверхности при взаимодействии с лазерным импульсом

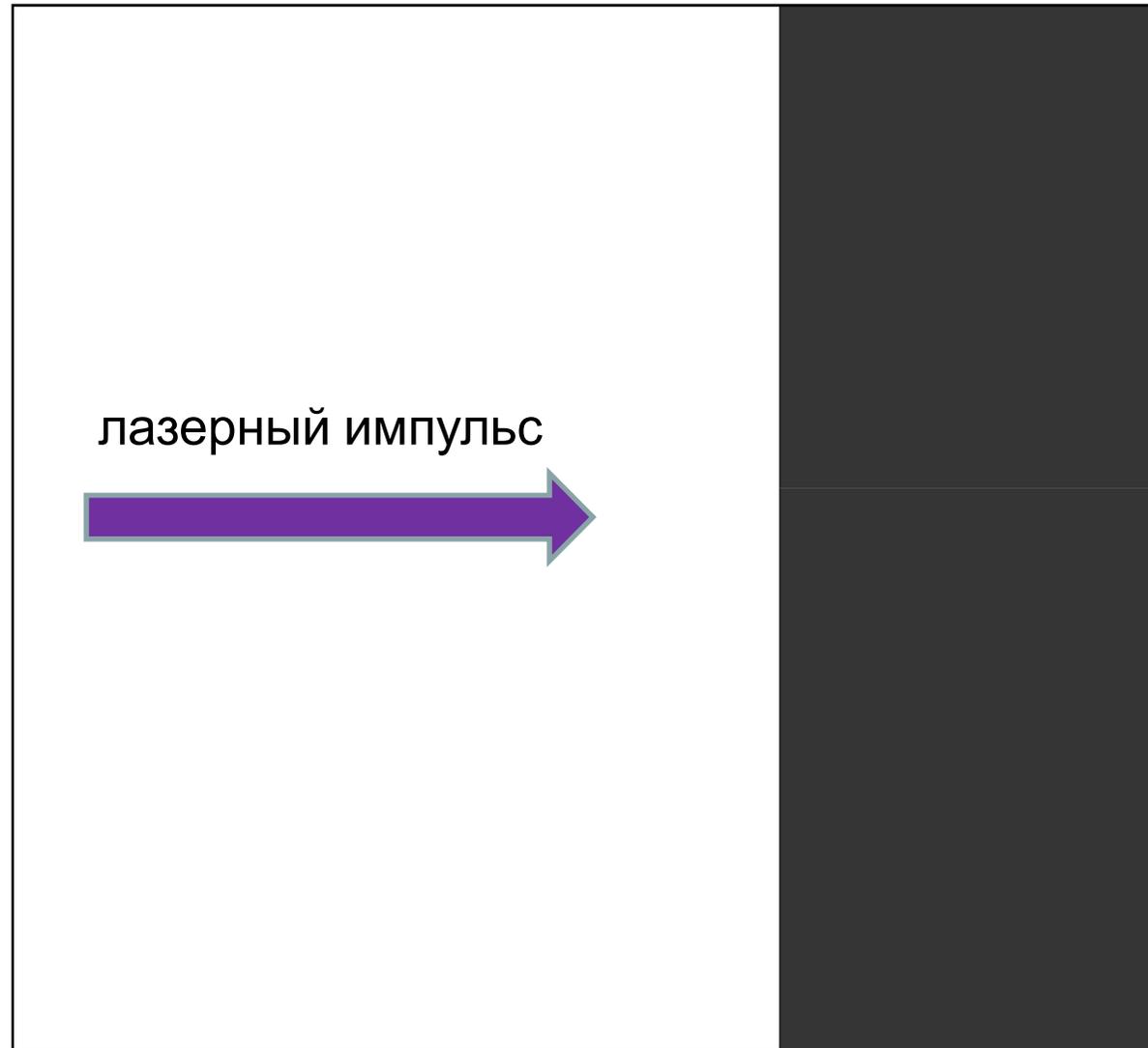


- [1] Starikov, Faenov, Pikuz, Skobelev, Fortov et al. // Appl. Phys. B: Lasers and Optics (2014)  
[2] Norman, Starikov, Stegailov, Saitov, Zhilyaev // Contrib. Plasma Phys. V. 53, P. 129 (2013)  
[3] Норман, Стариков, Стегайлов // ЖЭТФ, т. 141, стр. 910 (2012)

Зависимость глубины расплава/кратера от энерговклада  
для алюминия (1D-моделирование)



Атомистическое 2D-моделирование модификации поверхности  
лазерным импульсом  
*плавление*

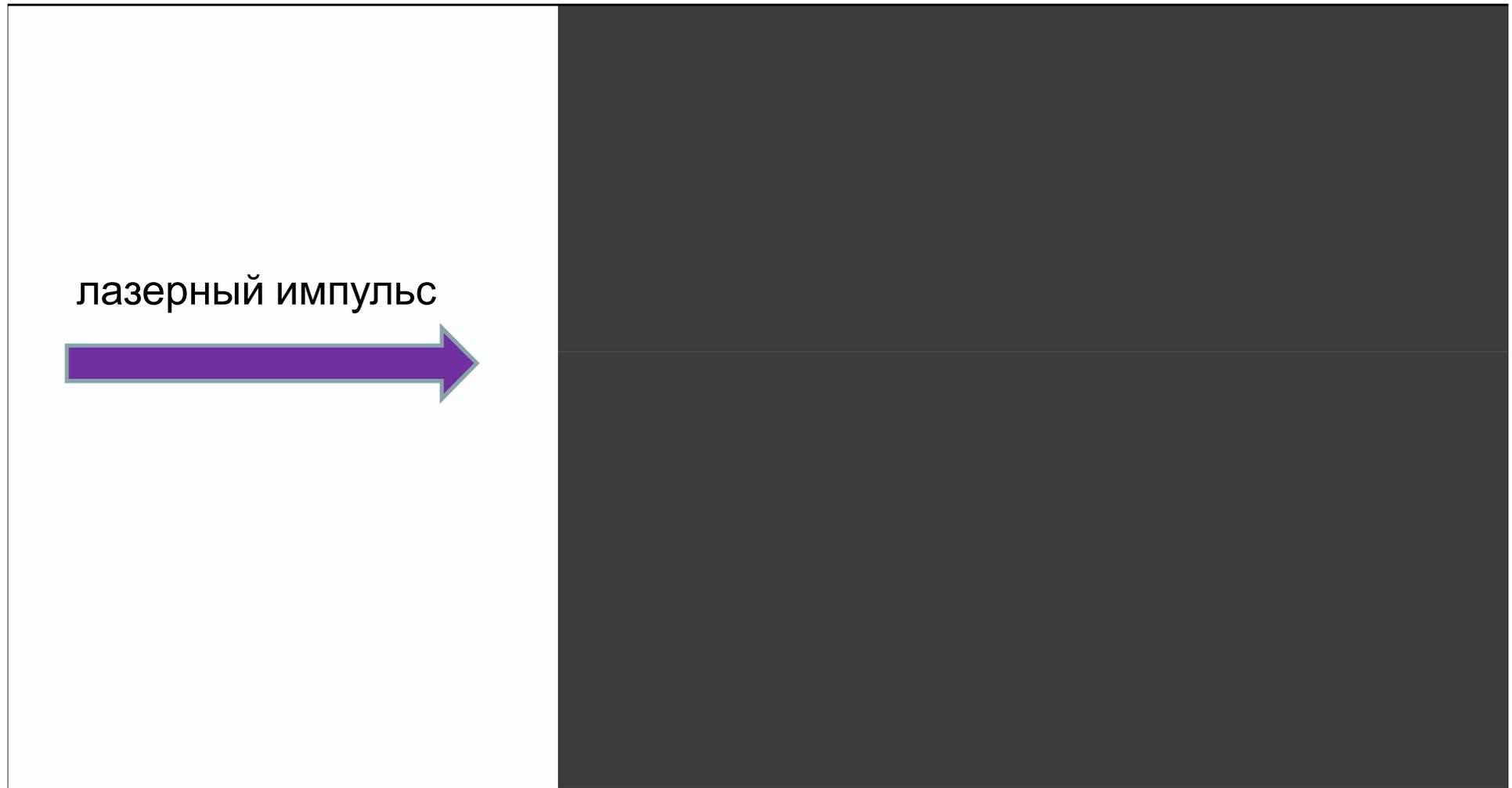


**1 мкм × 1 мкм × 10 нм**

**$F_0 = 30 \text{ мДж/см}^2$**

Атомистическое 2D-моделирование модификации поверхности  
лазерным импульсом

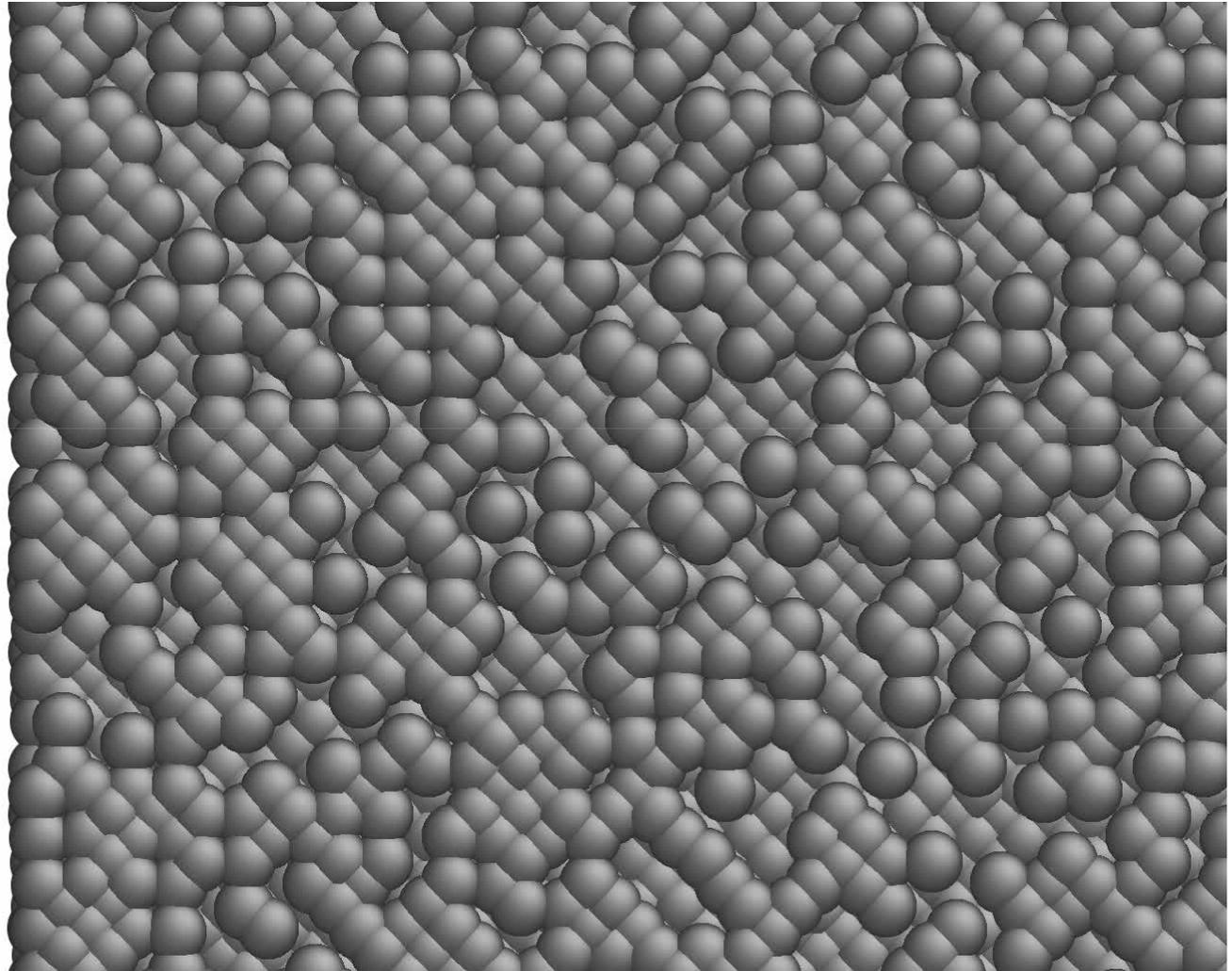
*абляция*



**1 мкм × 1 мкм × 10 нм**

**$F_0 = 55 \text{ мДж/см}^2$**

# Атомистическое 3D-моделирование модификации поверхности лазерным импульсом

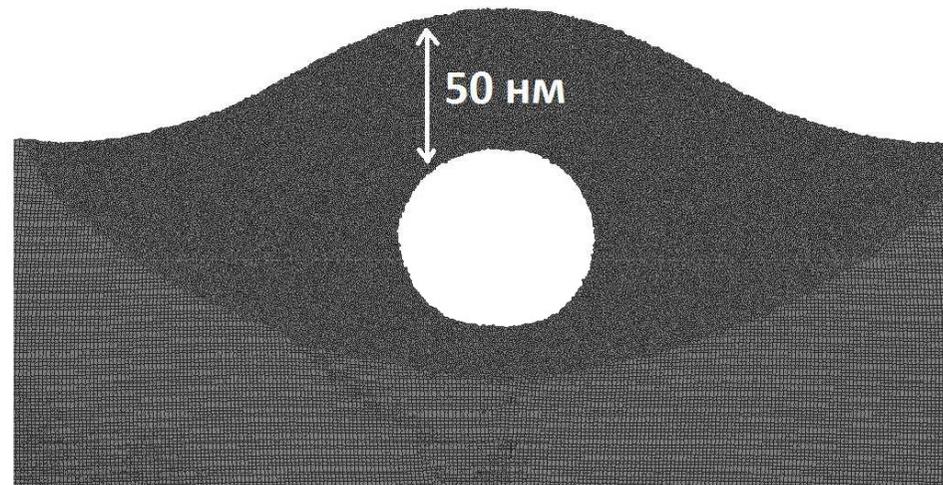
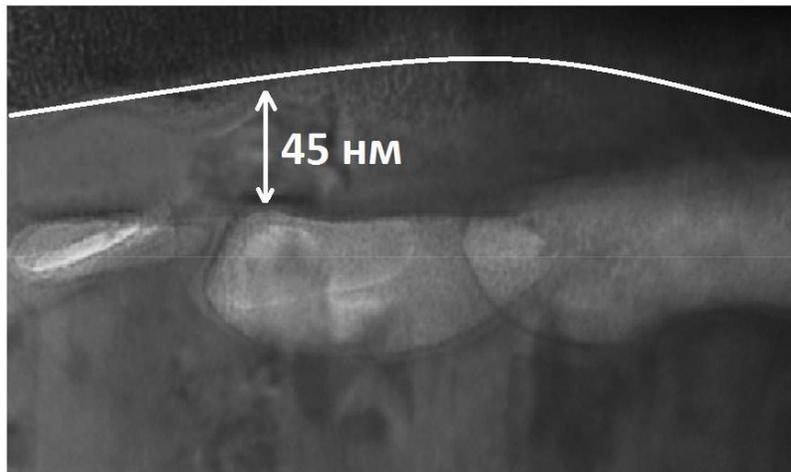


~  $10^9$  атомов

$$F_0 = 55 \text{ мДж/см}^2$$

# Атомистическое 3D-моделирование модификации поверхности лазерным импульсом

*нанополости в алюминии*



Изображение получено при  
использовании ТЕМ

[Ашитков, Иногамов, Жаховский, Эмиров,  
Агранат, Олейник, Анисимов, Фортов //  
Письма в ЖЭТФ (2012)]

Данная работа

## Выводы:

- Крупномасштабные расчеты лазерной абляции и модификации позволяют лучше понять механизм данного процесса. Предложены формулы описывающие размер области модификации. Высказана гипотеза о механизме формирования нанополостей вблизи поверхности.
- Показано, что переход к моделированию на больших масштабах даёт возможность исследовать новые коллективные явления атомов, недоступные для изучения на меньших масштабах.

**Оценка производительности  
кластера «Ломоносов»  
для задач классической молекулярной  
динамики**

# Масштабируемость задачи абляции (CPU) (weak scaling)

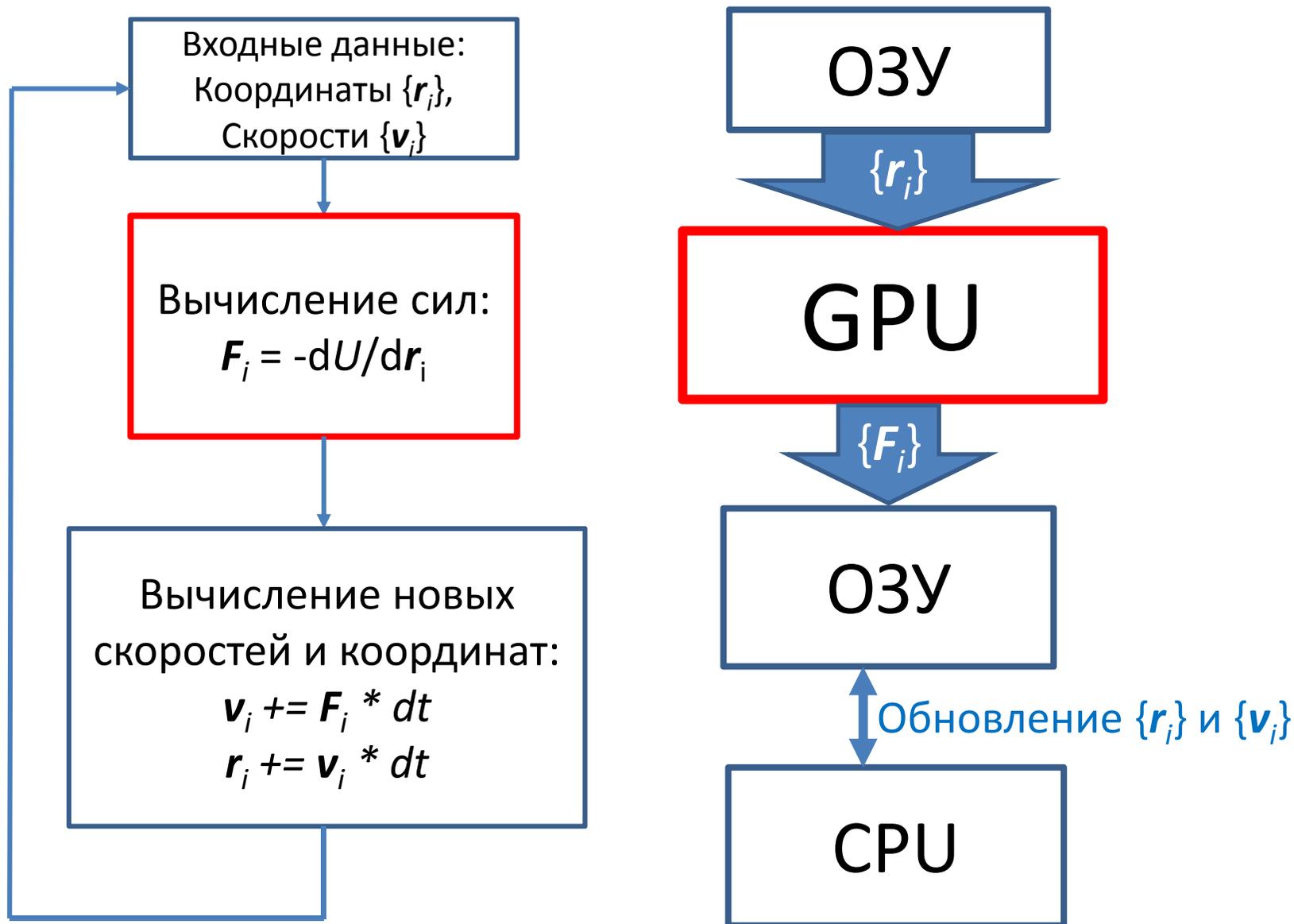


\* 1 узел = 8 ядер Intel  
Xeon

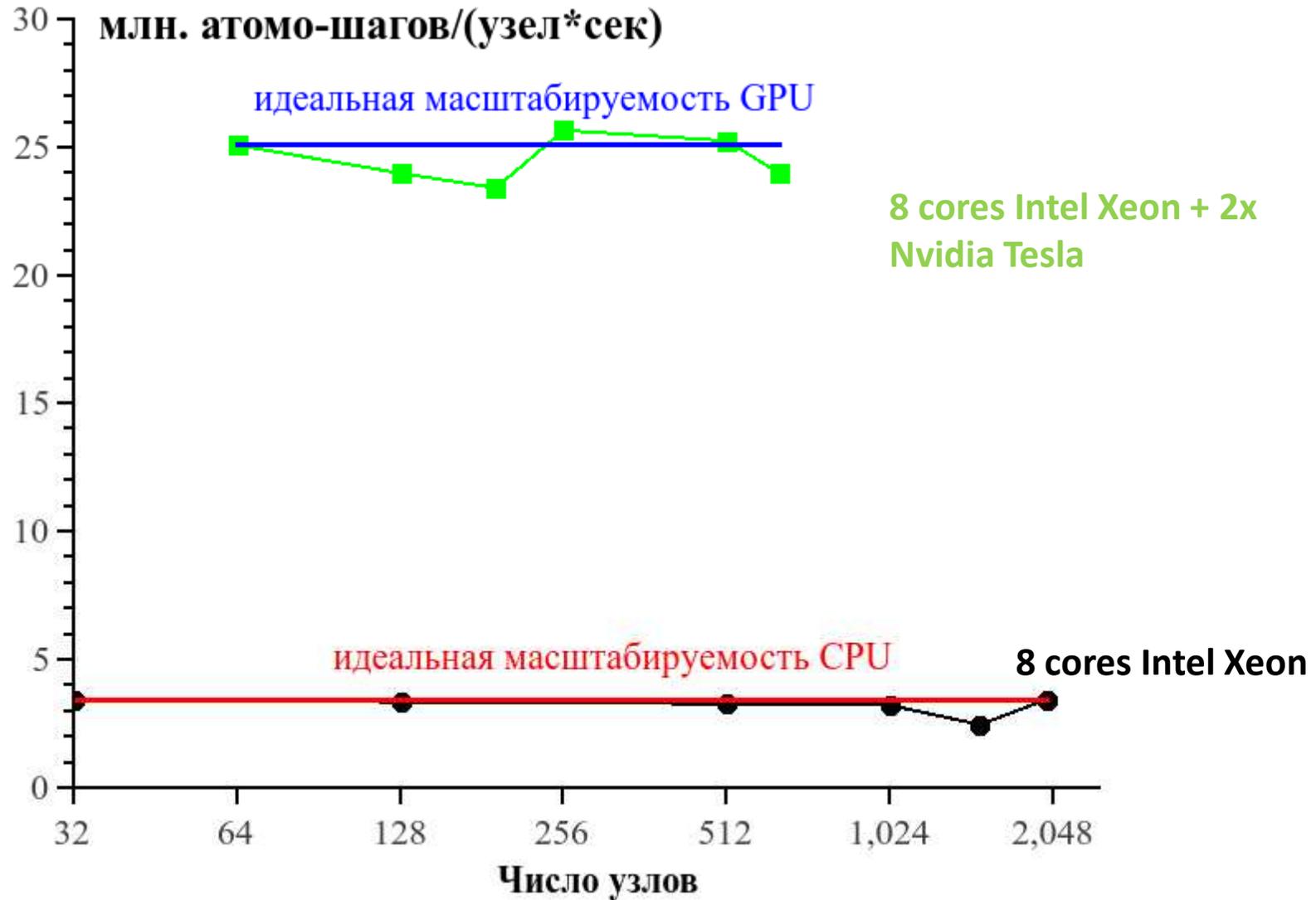
# GPU-ускорение молекулярной динамики



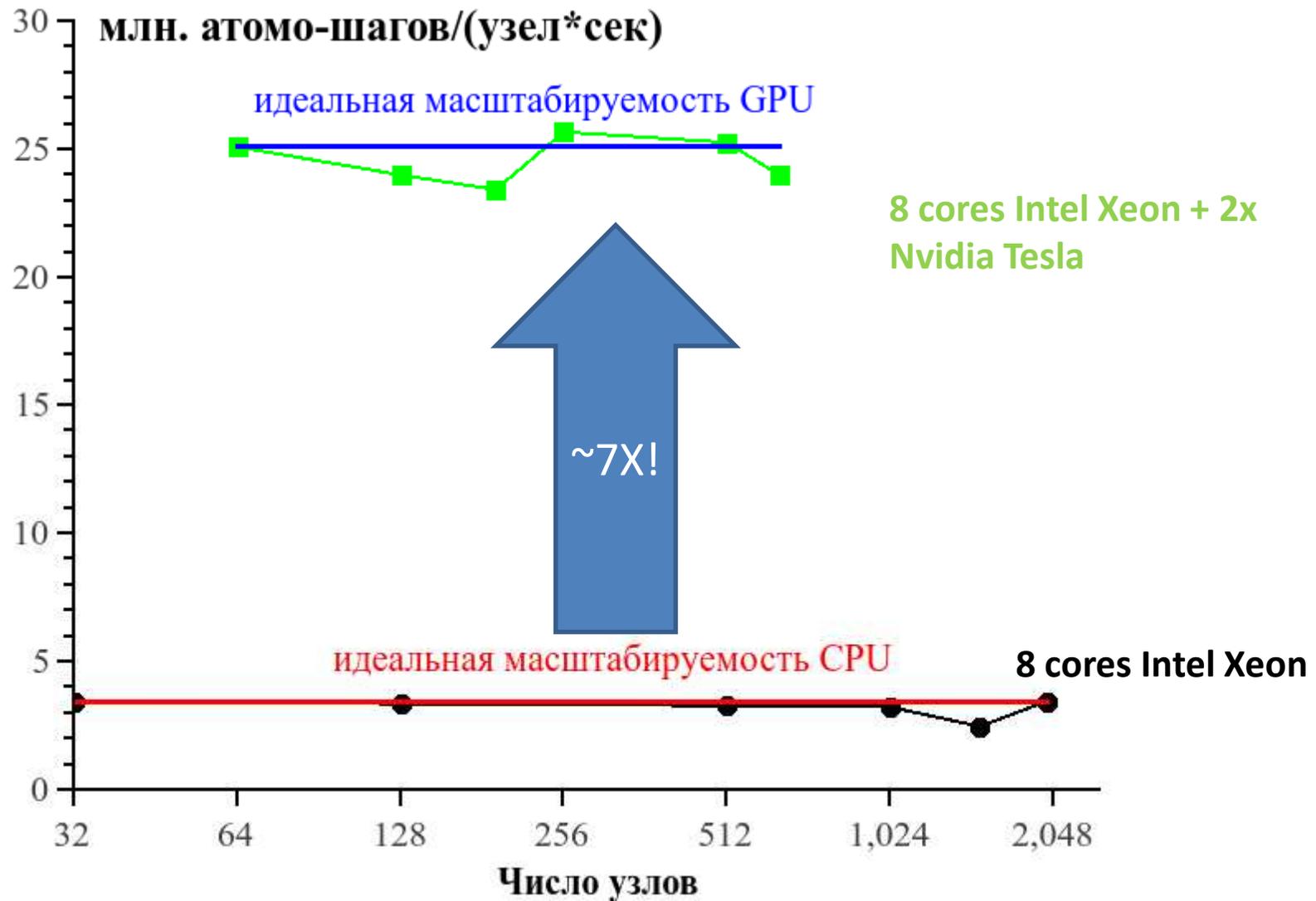
# GPU-ускорение молекулярной динамики



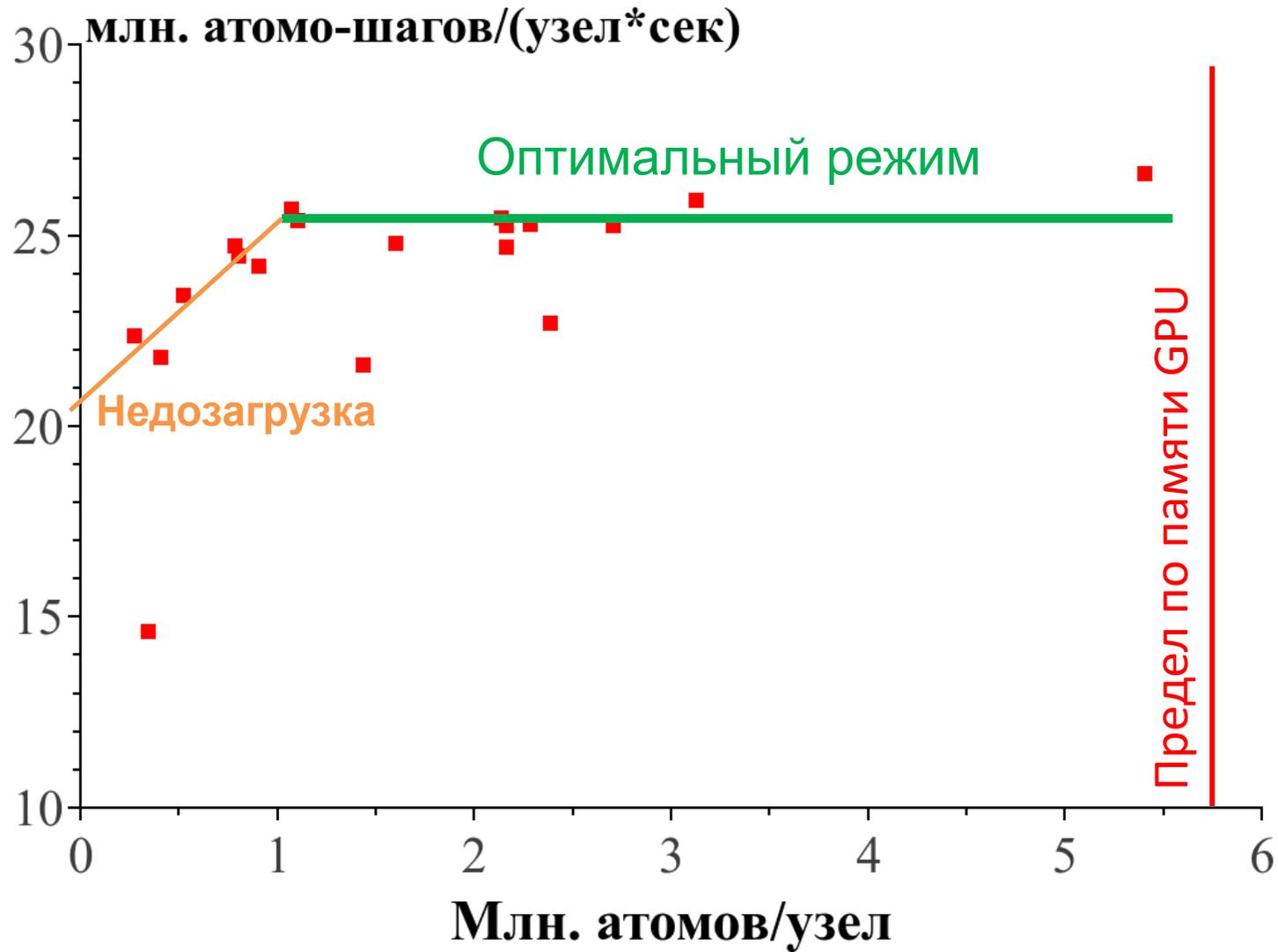
# Масштабируемость задачи абляции (CPU + GPU)



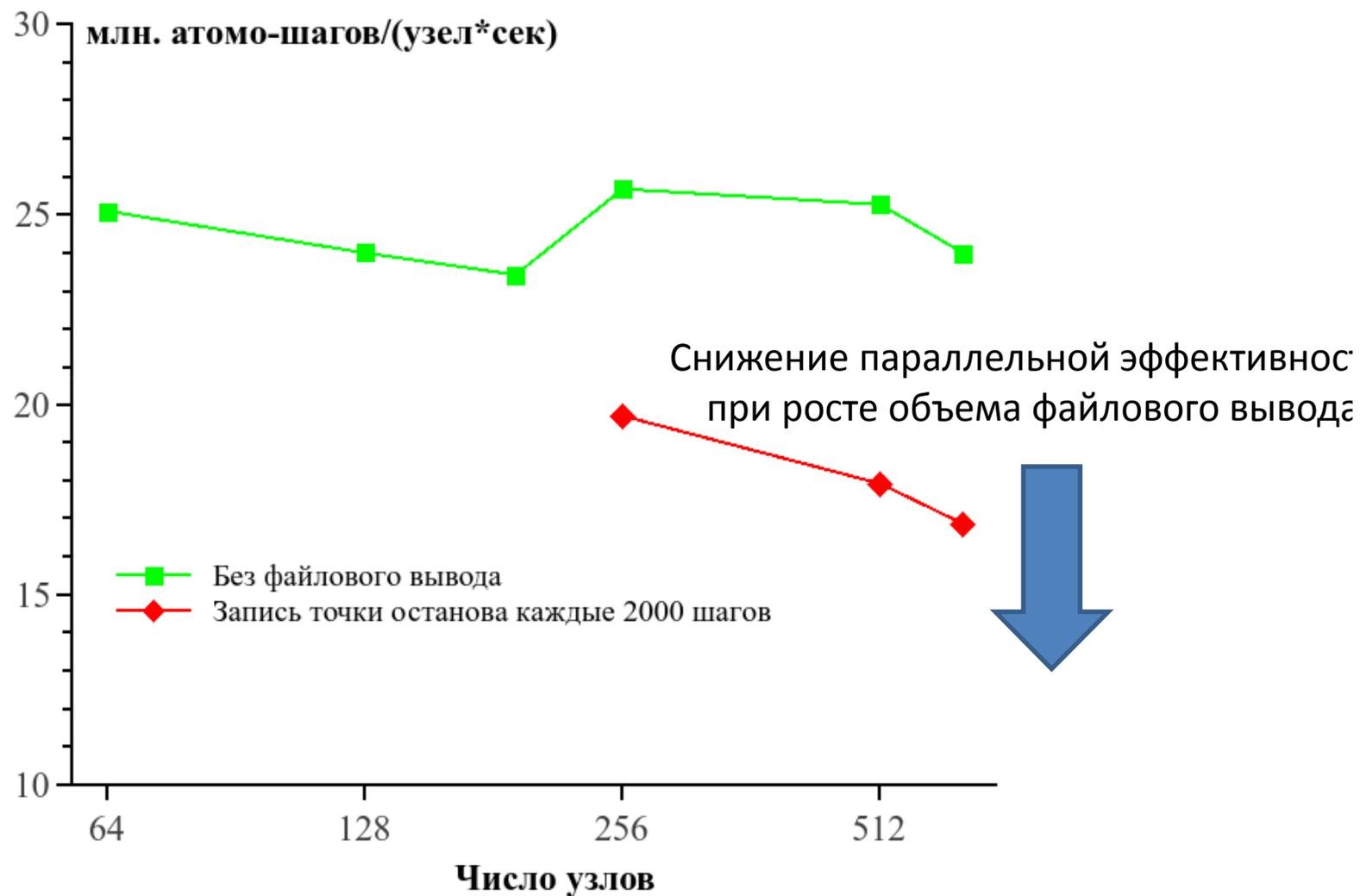
# Масштабируемость задачи абляции (CPU + GPU)



# Эффективность использования GPU



# Эффективность при выводе данных (GPU)



# Выводы:

## Преимущества видеоускорителей:

- Существенное ускорение МД расчетов
- На задачу требуется меньше узлов => меньше вероятность отказа

## Недостатки:

- Падение эффективности при недостаточной загрузке ускорителя
- Ограниченный объём памяти
- Меньше отношение pair time/comm time
- Не всегда есть оптимизированный под GPU код

# Общие проблемы крупномасштабного МД моделирования:

- Низкое время бесперебойного счета задачи – несколько часов
- Ошибки ввода-вывода на узлах, приводящие к остановке счета
- Невозможность прочитать в LAMMPS restart-файл при большом числе атомов (> 1 млрд.)

# Публикации по вычислительным аспектам работы

- Янилкин А.В., Жилаев П.А., Куксин А.Ю., Норман Г.Э., Писарев В.В., Стегайлов В.В. **Применение суперкомпьютеров для молекулярно-динамического моделирования процессов в конденсированных средах** // Вычислительные методы и программирование. 2010. Т. 11. С. 111-116. URL: [http://num-meth.srcc.msu.ru/zhurnal/tom\\_2010/v11r113.html](http://num-meth.srcc.msu.ru/zhurnal/tom_2010/v11r113.html)
- Morozov I.V., Kazennov A.M., Bystryi R.G., Norman G.E., Pisarev V.V., Stegailov V.V. **Molecular dynamics simulations of the relaxation processes in the condensed matter on GPUs** // Computer Physics Communications. 2011. V. 182. P. 1974–1978. <http://dx.doi.org/10.1016/j.cpc.2010.12.026>
- Жилаев П.А., Стегайлов В.В. **Ab initio молекулярная динамика: перспективы использования многопроцессорных и гибридных суперЭВМ** // Вычислительные методы и программирование. 2012. Т. 13. С. 37-45. URL: [http://num-meth.srcc.msu.ru/zhurnal/tom\\_2012/v13r206.html](http://num-meth.srcc.msu.ru/zhurnal/tom_2012/v13r206.html)
- Стегайлов В.В., Норман Г.Э. **Проблемы развития суперкомпьютерной отрасли в России: взгляд пользователя высокопроизводительных систем** // Программные системы: теория и приложения. 2014. Т. 5. № 1(19). С. 111–152. URL: [http://psta.psir.ru/read/psta2014\\_1\\_111-152.pdf](http://psta.psir.ru/read/psta2014_1_111-152.pdf)
- Куксин А.Ю., Ланкин А.В., Морозов И.В., Норман Г.Э., Орехов Н.Д., Писарев В.В., Смирнов Г.С., Стариков С.В., Стегайлов В.В., Тимофеев А.В., **ЗАЧЕМ и КАКИЕ нужны суперкомпьютеры эксафлопсного класса? Предсказательное моделирование свойств и многомасштабных процессов в материаловедении** // Программные системы: теория и приложения. 2014. Т. 5. № 1(19). С. 175. (в печати)

## Публикации по лазерной абляции

- Стариков С.В., Стегайлов В.В., Норман Г.Э., Фортов В.Е., Ишино М., Танака М., Хасегава Н., Нишикино М., Охба Т., Каихори Т., Очи Е., Имазоно Т., Кавачи Т., Тамотсу С., Пикуз Т.А., Скобелев И.Ю., Фаенов А.Я. **Лазерная абляция золота: эксперимент и атомистическое моделирование** // Письма в ЖЭТФ. 2011. Т. 93. N. 11. С. 719-725. [http://www.jetpletters.ac.ru/ps/1934/article\\_29313.shtml](http://www.jetpletters.ac.ru/ps/1934/article_29313.shtml)
- Норман Г.Э., Стариков С.В., Стегайлов В.В. **Атомистическое моделирование лазерной абляции золота: эффект релаксации давления** // ЖЭТФ. 2012. Т. 141. N. 5. С. 910-918. <http://dx.doi.org/10.1134/S1063776112040115>
- Norman G.E., Starikov S.V., Stegailov V.V., Fortov V.E. et al. **Nanomodification of gold surface by picosecond soft x-ray laser pulse** // J. Appl. Phys. 2012. V. 112. P. 013104. URL: <http://dx.doi.org/10.1063/1.4731752>
- Starikov S.V., Faenov A.Ya., Pikuz T.A., Skobelev I.Yu., Fortov V.E. **Soft picosecond X-ray laser nanomodification of gold and aluminum surfaces** // Appl. Phys. B. 2014. URL: <http://dx.doi.org/10.1007/s00340-014-5789-y>