

Моделирование “из первых принципов” молекулярных переключателей на суперкомпьютерах.

Лекция 9

факультет ВМК МГУ

профессор А.М.Попов

Цели проекта.

- **Нанотехнологии.** Суперкомпьютерное моделирование как часть технологического процесса.
- Разработка программ для моделирования молекулярных переключателей “из первых принципов” для задач нанoeлектроники. Расчет новых молекулярных переключателей, нанопамяти, транспортных процессов в молекулярных наноструктурах.
- Разработка **многомасштабной самосогласованной модели** на основе квантовой молекулярной динамики Эрнфеста. Нелокальные эффекты переключения. Влияние подложки.
- **Эффективность суперкомпьютера.**
- Анализ взаимосвязей между структурой параллельного алгоритма, требованиями к памяти и архитектурами многопроцессорной вычислительной системы с общей и распределенной памятью.
- **Создание параллельной системы визуализации результатов научных исследований в реальном времени на Blue Gene/P МГУ.**

Молекулярные переключатели.

- Производительность компьютера пропорциональна количеству транзисторов на единице площади интегральной схемы. N/S . Сейчас $N \sim 10^9$ - это предел.
- Толщина изолирующего слоя оксида кремния между затвором и проводящим слоем – критический элемент. Предел ~ 2 нм. Меньшие размеры невозможны на базе стандартной техники фотолитографии. Необходимы принципиально новые технологии.
- Актуальность технологических исследований проводимых корпорациями, создающими суперкомпьютеры (IBM) обусловлена физическим пределом кремниевых технологий.

9

- Молекулярные переключатели. Размеры на два порядка меньше кремниевых. Размер транзистора порядка одного нанометра. В миллион раз больше транзисторов. Время отклика уменьшится до фемтосекунд (на шесть порядков) — время протекания элементарной стадии химической реакции.
- Эффективность молекулярного компьютера может оказаться в миллиарды раз выше, чем современного кремниевого.
- Моделирование процессов такой высокой сложности на суперкомпьютерах с массовым параллелизмом способно предсказывать поведение сложных молекулярных наносистем.

Задачи нанотехнологий.

Молекулярная электроника.

- **Построение переключателей на молекулярном уровне.**
- **Сверхплотная память 100Гб на 1 кв. см x100**
- **Молекулярный компьютер**
- **Молекулярный ассемблер. Самосборка.**

Г.Бенниг,Г.Рорер.(IBM). Нобелевская премия 1986г. Сканирующий туннельный микроскоп – видит атомы. Разрешающая способность – 0.001nm

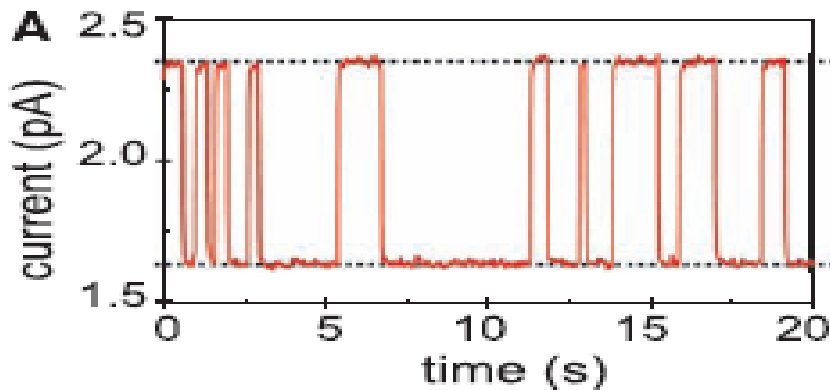
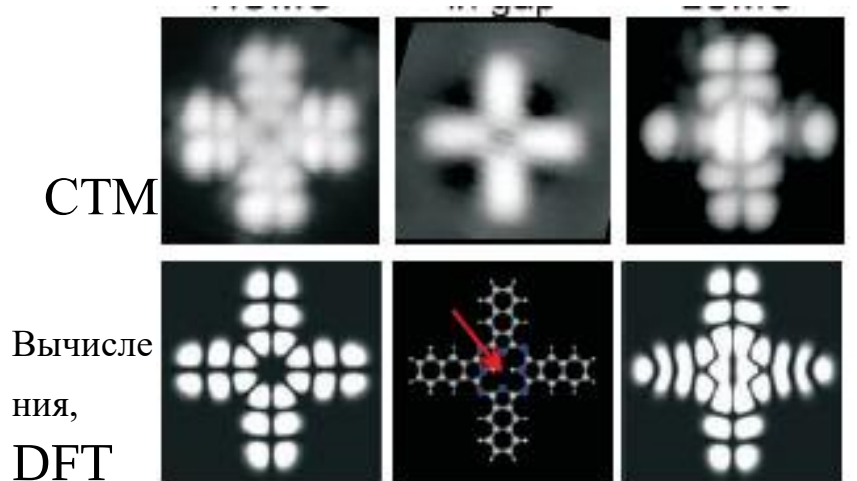
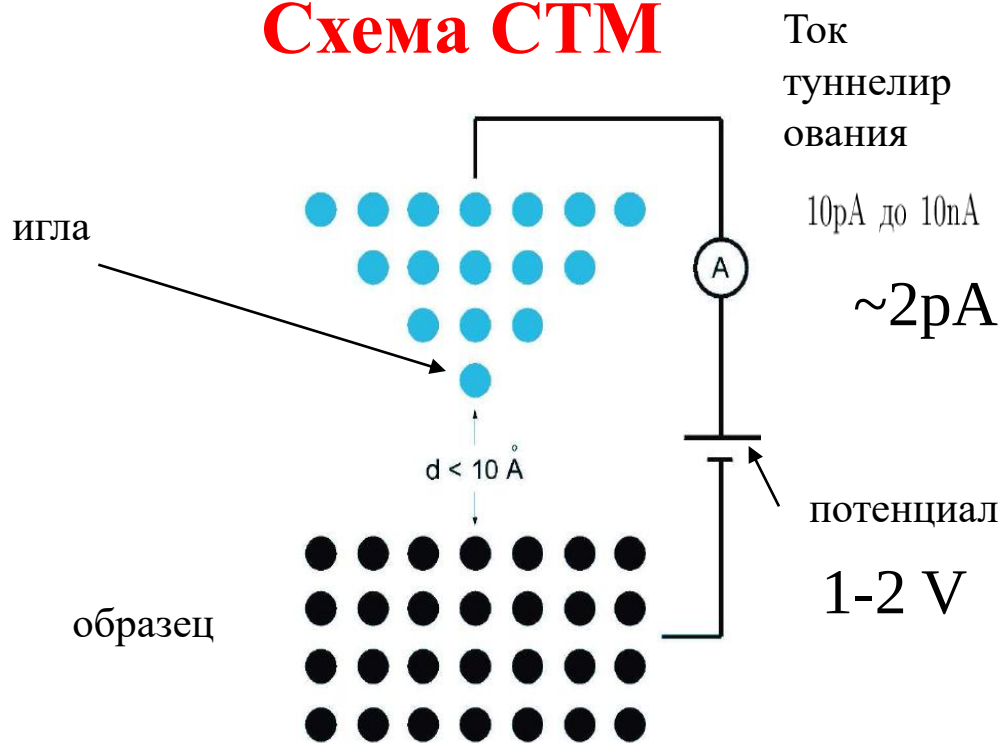
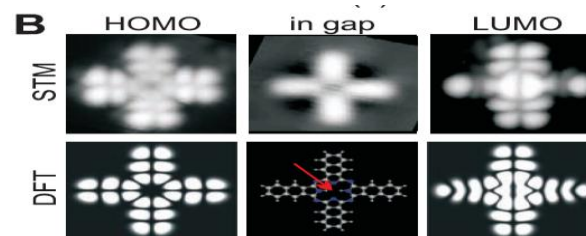
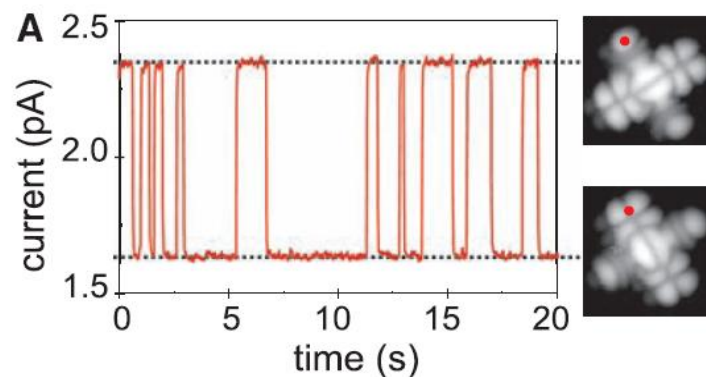
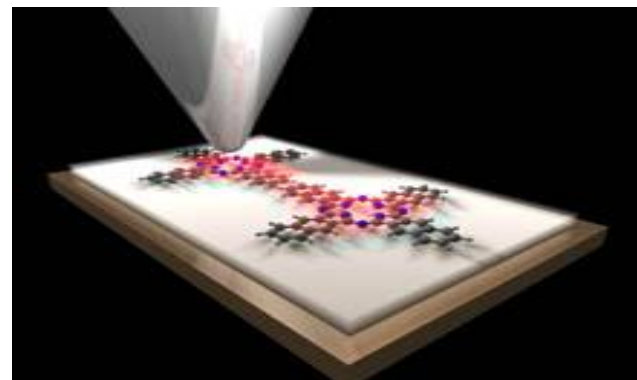


Схема СТМ



Моделирование нового типа молекулярных переключателей

- Эксперимент в Цюрихе. Инициированная током реакция водородной таутомеризации молекулы нафталоцианина и переключение проводимости. Регистрация тока туннелирования STM.
- Новый молекулярный переключатель. **Переключение проводимости без изменения формы молекулы** в зависимости от положения двух внутренних атомов водорода в центральной области молекулы.
- Реакция эквивалентна повороту молекулы на 90 град. и обуславливает изменение измеряемого тока туннелирования.

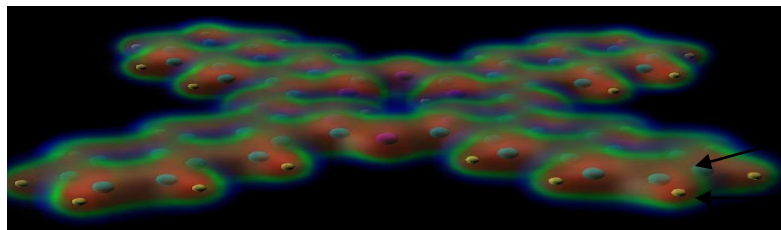


Peter Liljeroth, et al. Science 317, 1203 (2007)

LUMO нафталоцианина может иметь две ориентации

Вычислительные модели “из первых принципов” для наносистем

- Молекулярная динамика из первых принципов FPMD – метод моделирования наносистем. Нет подгоночных параметров. **Проблема многих тел.**



электроны

ядра

Только
кулоновские
взаимодейст
вия

- Вычислительная проблема. Наномасштаб - мезомасштаб (атом – спл.среда)

Куб полиэтилена - размером 100nm – $64 \cdot 10^6$ атомов

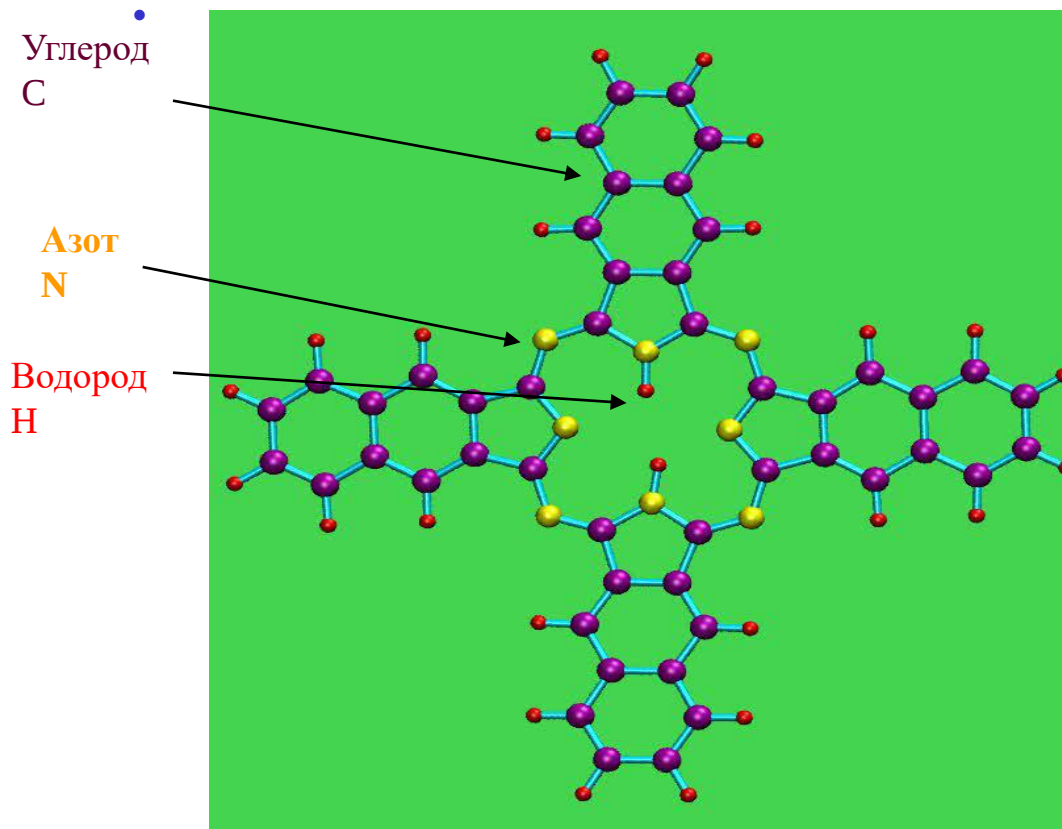
Куб железа - в 1nm – 16000 атомов

- Квантовомеханическое описание электронов и классическое описание движения атомных ядер
- Подход требует использования суперкомпьютеров и эффективных параллельных вычислений для описания больших молекулярных систем ~ 25000 атомов.
- Масштабируемость FPMD кодов.

Вычислительные квантовомеханические модели “из первых принципов” для наносистем

Почему параллельные вычисления принципиально важны?

Пример моделирования молекулярного переключателя. Смена положения атомов водорода в полости молекулы нафталоцианина не меняет геометрию молекулы. Расчет на Blue Gene/P (МГУ)



Пример небольшой системы молекулярного переключателя: 82 атома, 129 орбиталей, 540000 пл. волн для каждой в. функции, шаг = 0.1 fs, сетка 320x320x192

Время вычислений

512 процессоров BG/P
– 20 часов

Однопроцессорный компьютер – 1 год!!

Алгоритмы FPMD

- Классическая МД для ядер

$$M_I \ddot{\mathbf{R}}_I(t) = - \frac{\partial E^{(0)}(\{\psi_n^{(0)}\}, \{\mathbf{R}_I\})}{\partial \mathbf{R}_I}$$

- Расчет электронной структуры. Уравнение Шредингера для основного состояния

$$H_e[\{\psi_n^{(0)}\}, \{\mathbf{R}_I\}] \psi^{(0)} = E^{(0)} \psi^{(0)}$$

Уравнения CPMD. Вычисления на BG/P в МГУ

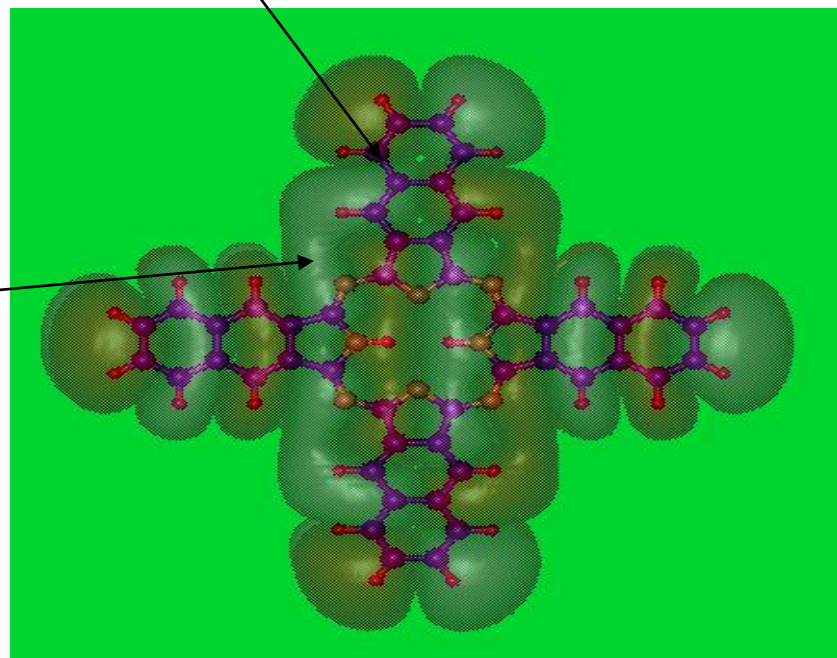
$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} E_{KS}[\{\psi_n\}, \{\mathbf{R}_I\}]$$

ядра

$$\mu_n \ddot{\psi}_n(t) = -\frac{\delta}{\delta \psi_n^*} E_{KS}[\{\psi_n\}, \{\mathbf{R}_I\}] + \sum_{n'} \Lambda_{nn'} \psi_{n'}$$

НОМО орбиталь

$$\rho(\mathbf{r}) = \sum_n^{N_e} |\psi_n(\mathbf{r})|^2$$



Расчет электронной структуры. DFT.

- **Уравнения Кона-Шэма. Нелинейная задача на собственные значения.**

$$\left(-\frac{1}{2}\Delta + U_{eff}(\rho, \mathbf{r})\right)\psi_n(\mathbf{r}) = \epsilon_n\psi_n(\mathbf{r})$$

- **Электронная плотность**

$$\rho(\mathbf{r}) = \sum_n^{N_e} |\psi_n(\mathbf{r})|^2$$

- **Эффективный потенциал**

$$U_{eff}(\mathbf{r}) = U_{ion}(\mathbf{r}) + \int \frac{\rho(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s} + U_{XC}(\rho\mathbf{r}) + U_I$$

Координационная переменная реакции изомеризации.

$$s_1(H_1, H_2, N_1, N_2) = \frac{1}{2} \left\{ \frac{1 - \left(\frac{R_{H_1N_1}}{R_0}\right)^6}{1 - \left(\frac{R_{H_1N_1}}{R_0}\right)^{12}} + \frac{1 - \left(\frac{R_{H_1N_2}}{R_0}\right)^6}{1 - \left(\frac{R_{H_1N_2}}{R_0}\right)^{12}} + \frac{1 - \left(\frac{R_{H_2N_1}}{R_0}\right)^6}{1 - \left(\frac{R_{H_2N_1}}{R_0}\right)^{12}} + \frac{1 - \left(\frac{R_{H_2N_2}}{R_0}\right)^6}{1 - \left(\frac{R_{H_2N_2}}{R_0}\right)^{12}} \right\}$$

$R_{H_iN_j}$ -расстояния между i -ым атомом водорода Н и j -ым атомом азота N

$R_0 = 1.05 \text{ \AA}$

$$s_2 = s_1(H_1, H_2, N_1, N_2) - s_1(H_1, H_2, N_3, N_4)$$

Путь химической реакции в терминах координационных переменных

Внутренняя координата реакции – наикратчайший путь на поверхности потенциальной энергии, который соединяет реактант и продукт.

Свободная энергия Гиббса вдоль координационной переменной. Метод Метадинамики.

Метадинамика – метод вычисления поверхности свободной энергии вдоль коорд. переменной (ых) $s(t)$ (длин связей, торсионных углов) связанной с коллективной переменной $S(t)$ ядер.

Свободная энергия $F(s) = -k_B T \ln[P(s)]$

k_B -константа Больцмана, T - температура, P - функция -распределения

$$P(s) = \frac{1}{Z} \int d\mathbf{r} \cdot e^{-\frac{U(\mathbf{r})}{k_B T}} \delta(s(\mathbf{r}) - s)$$

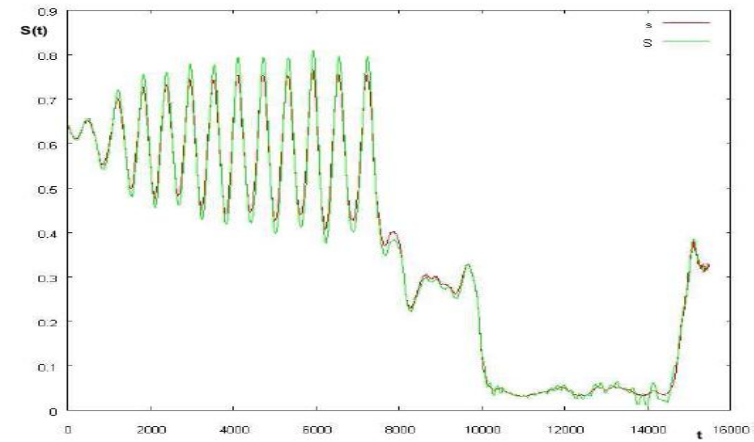
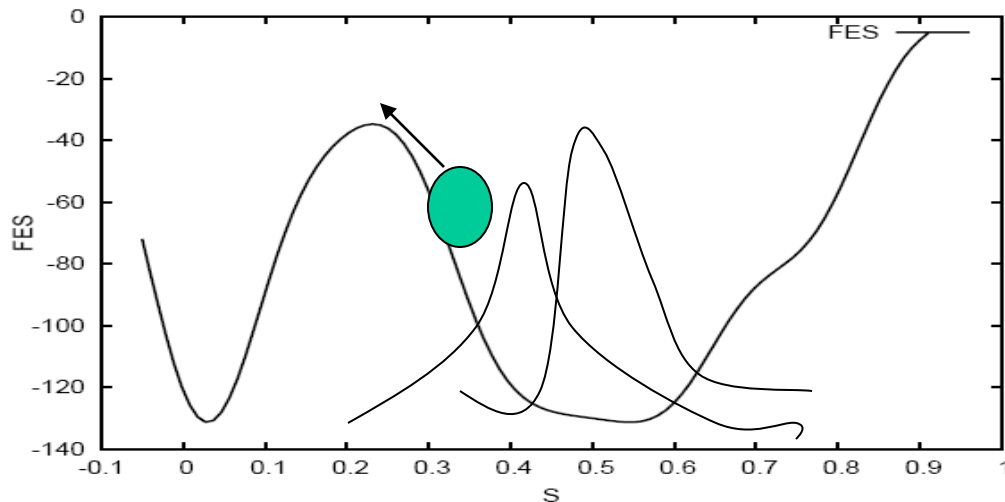
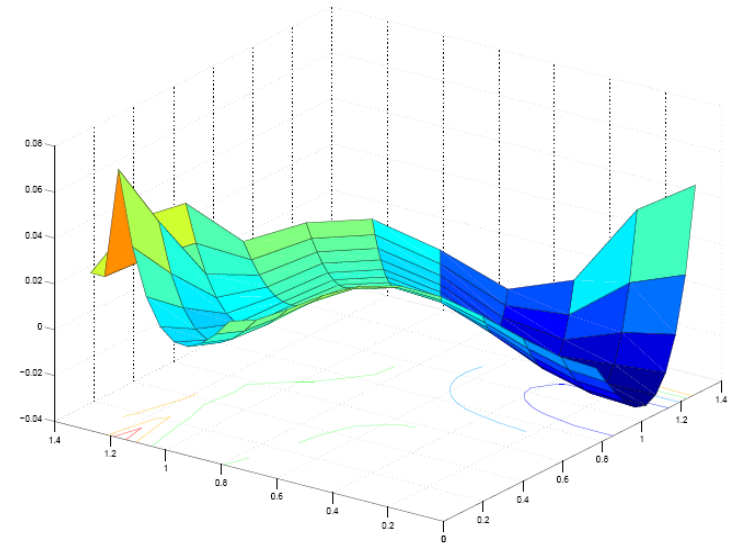
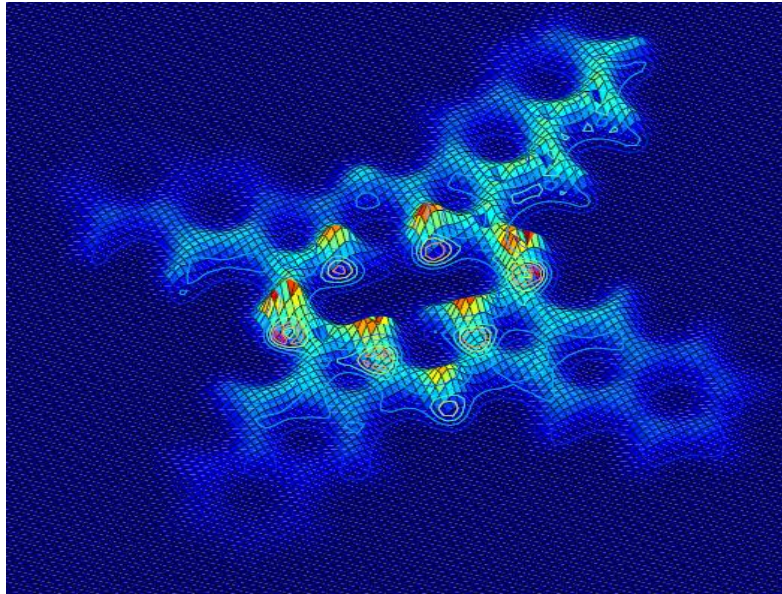
Искусственный потенциал $V(s,t)$ добавляется в систему

$$V(s, t) = - \sum_{t_i < t} H \cdot \exp\left[-\frac{(s(t) - s^i)^2}{2W^2}\right]$$

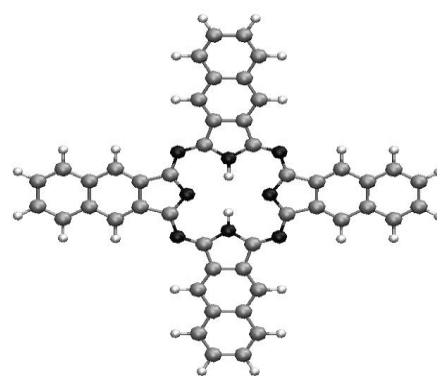
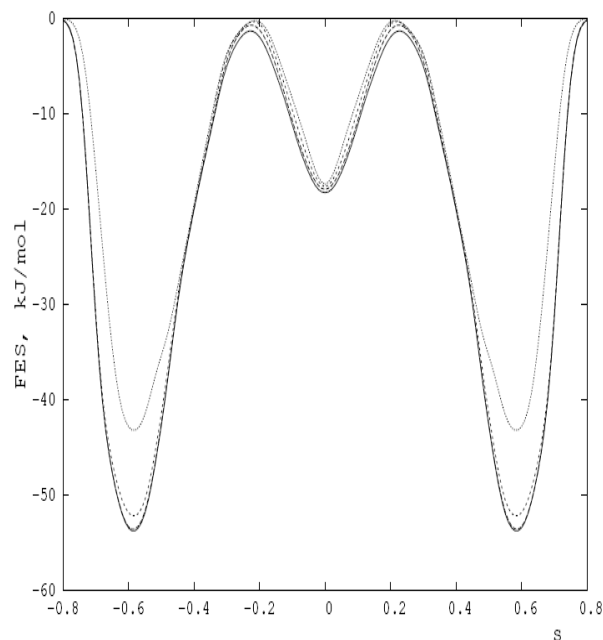
Потенциальные холмы добавляются в зависящий от истории потенциал $V(s,t)$ с

$$\tau_{MTD} \gg \tau_{MD}$$

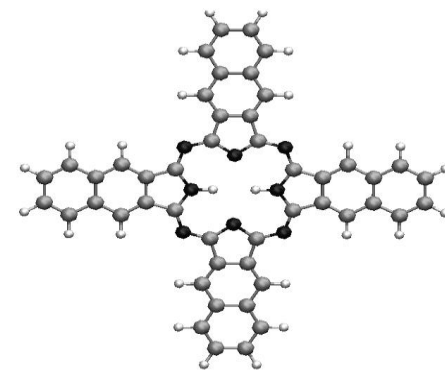
Параллельная метадинамика для моделирования барьера переключения. Добавление гауссианов для вычисления FES. 512 $\rightarrow$ 2048 ускорение = 4.



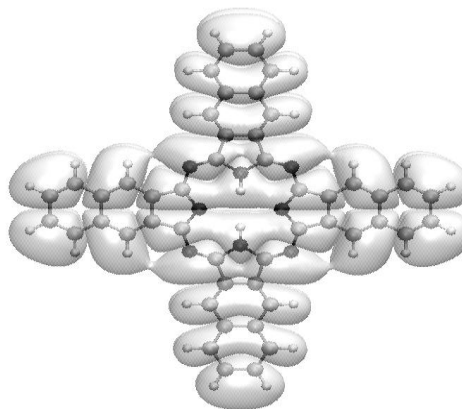
Код CPMD. Вычисления на BG/P в МГУ. Профиль свободной энергии вдоль координационной переменной. Метод метадинамики. Многократные случайные блуждания.



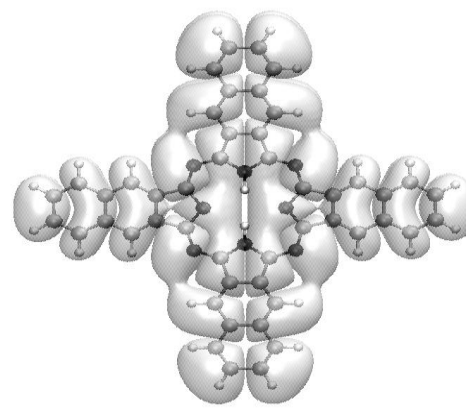
а



б



в



г

Рис. 2: Вычисленный профиль поверхность свободной энергии вдоль координатной переменной "s". Сходимость достигнута в результате четырех последовательных расчетов с параметрами : 1) $H = 3k_B T$; $\Delta s = 0.03$ 2) $H = 1.5k_B T$; 3) $H = 0.375k_B T$

Разработана **многомасштабная модель** молекулярного переключения. Нелокальность. FES из CPMD. Переключения на основе возбужденных орбиталей из МД Эрнфеста.

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum_m c_m(t) \psi_m(\mathbf{r}, \mathbf{R}) \quad \mathbf{R} \rightarrow s$$

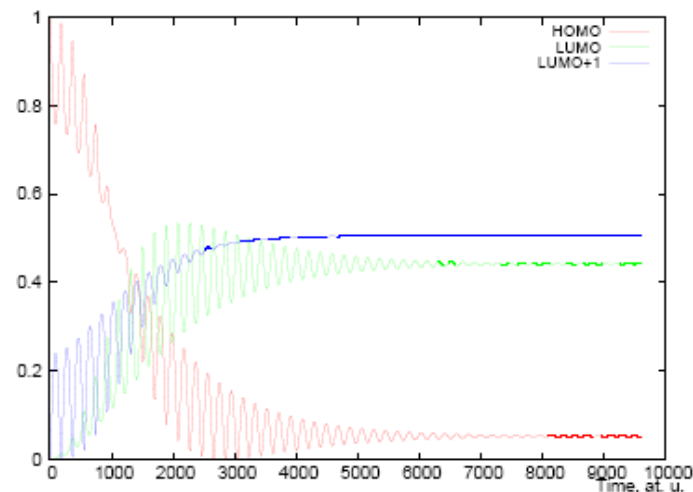
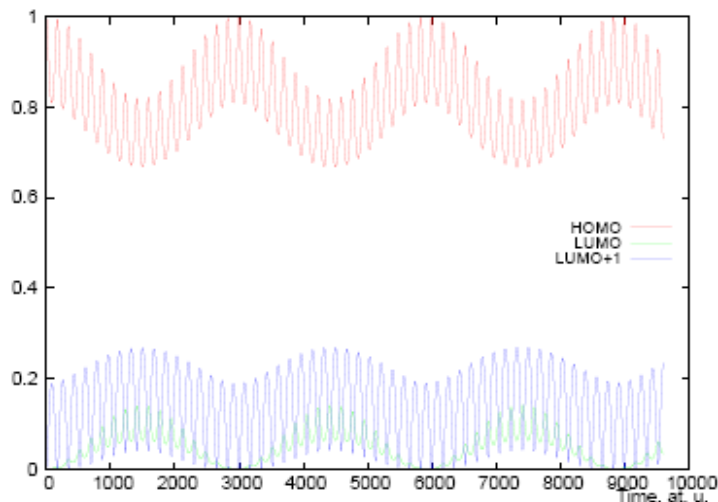
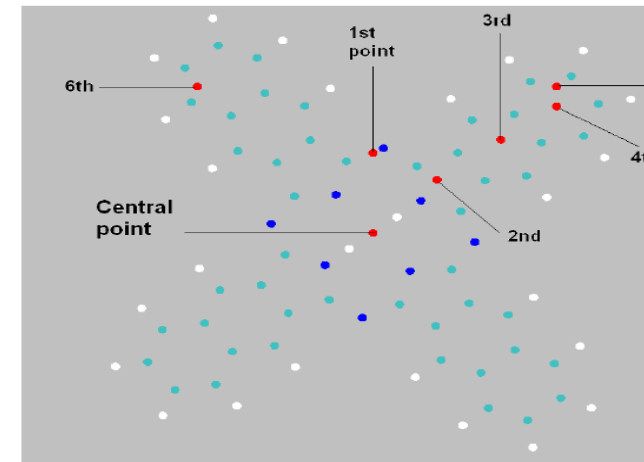
$$V_{mk}^{ext} = \langle \psi_m | V^{ext} | \psi_k \rangle$$

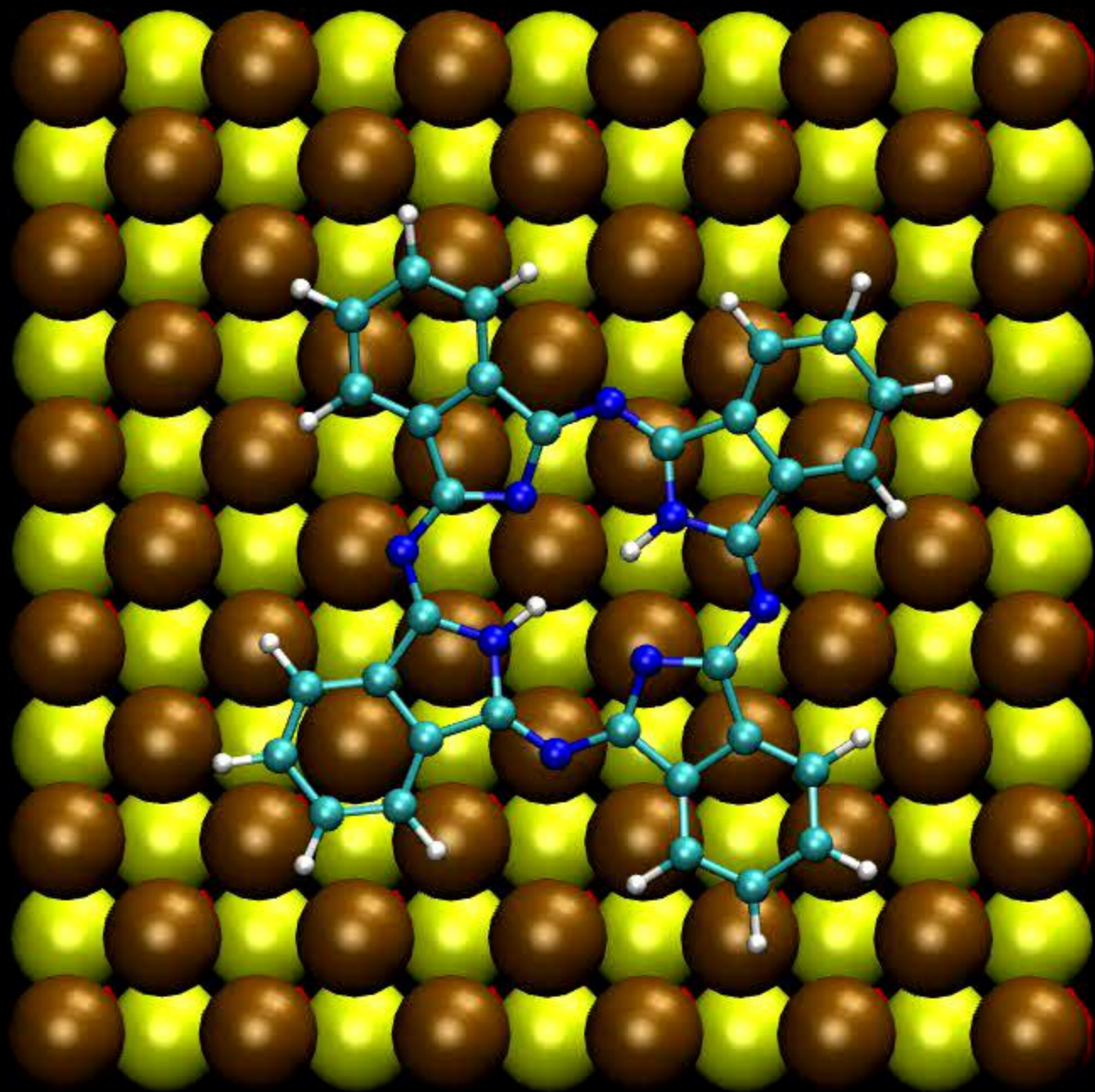
$$\mu \frac{d^2 s}{dt^2} = - \frac{\partial U_F}{\partial s} - \sum_{m > m_0} |c_m|^2 \frac{\partial E_m}{\partial s}(s) - \sum_{m > m_0} \sum_{k \neq m} c_m(t) c_k(t) (E_m - E_k) d_{mk}(s)$$

$$i \frac{dc_m(t)}{dt} = c_m(t) E_m(s) - i \sum_k c_k(t) \frac{ds}{dt} d_{mk}(s) + \sum_k V_{mk}^{ext} c_k,$$

$$V_0 = 1.4 \text{ eV},$$

$$V_0 = 1.5 \text{ eV}.$$





Расчет электронной структуры. **DFT.**

Основная вычислительная сложность.

Итерации по плотности электронов.

Вычислительная сложность $\sim N^3$, N – электронов.

- Уравнения Кона-Шэма. Нелинейная задача на собственные значения.

$$\left(-\frac{1}{2}\Delta + U_{eff}(\rho, \mathbf{r})\right)\psi_n(\mathbf{r}) = \epsilon_n\psi_n(\mathbf{r})$$

$$\psi_n(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{K}} c_{n,\mathbf{K}} e^{i\mathbf{K}\mathbf{r}}$$

- Электронная плотность

$$\rho(\mathbf{r}) = \sum_n^{N_e} |\psi_n(\mathbf{r})|^2$$

Прямое 3D FFT

$$c_{n,\mathbf{K}} \rightarrow \mathbf{C}$$

- Эффективный потенциал

$$U_{eff}(\mathbf{r}) = U_{ion}(\mathbf{r}) + \int \frac{\rho(\mathbf{s})}{|\mathbf{r} - \mathbf{s}|} d\mathbf{s} + U_{XC}(\rho\mathbf{r}) + U_I$$

$$\hat{H}\mathbf{C} = \epsilon\mathbf{C}$$

Обратное 3D FFT

$$U_{eff}(\rho, \mathbf{r})$$

Параллельные стратегии. Несколько уровней распараллеливания

- **Крупноблочное распараллеливание** на распределенной памяти **MPI** – распределение коэффициентов волновых функций для всех электронных состояний на все процессоры.

3dFFT Реальное пространство  К-пространство

Распределение данных минимизирует число передач с поддержанием загрузки в обоих пространствах.

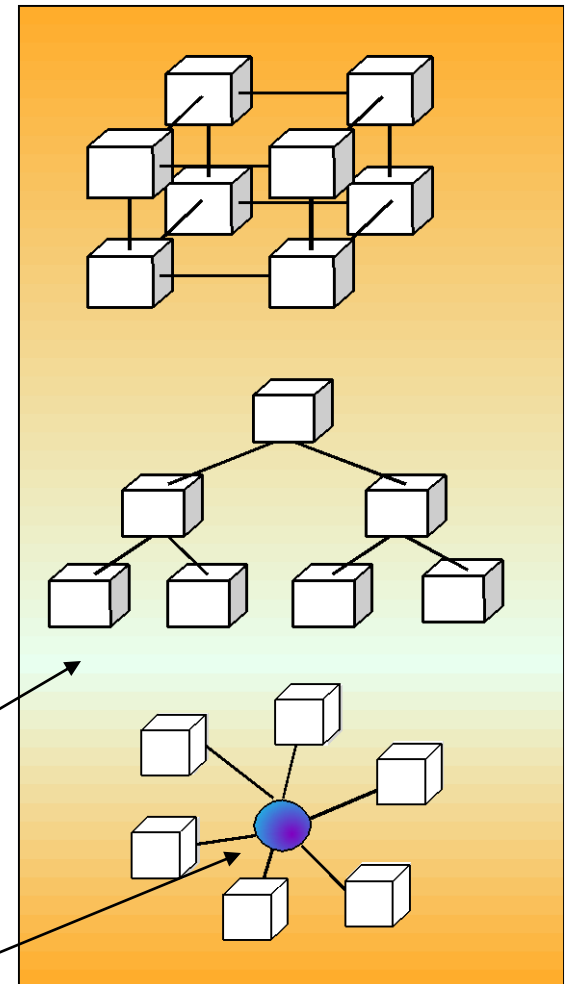
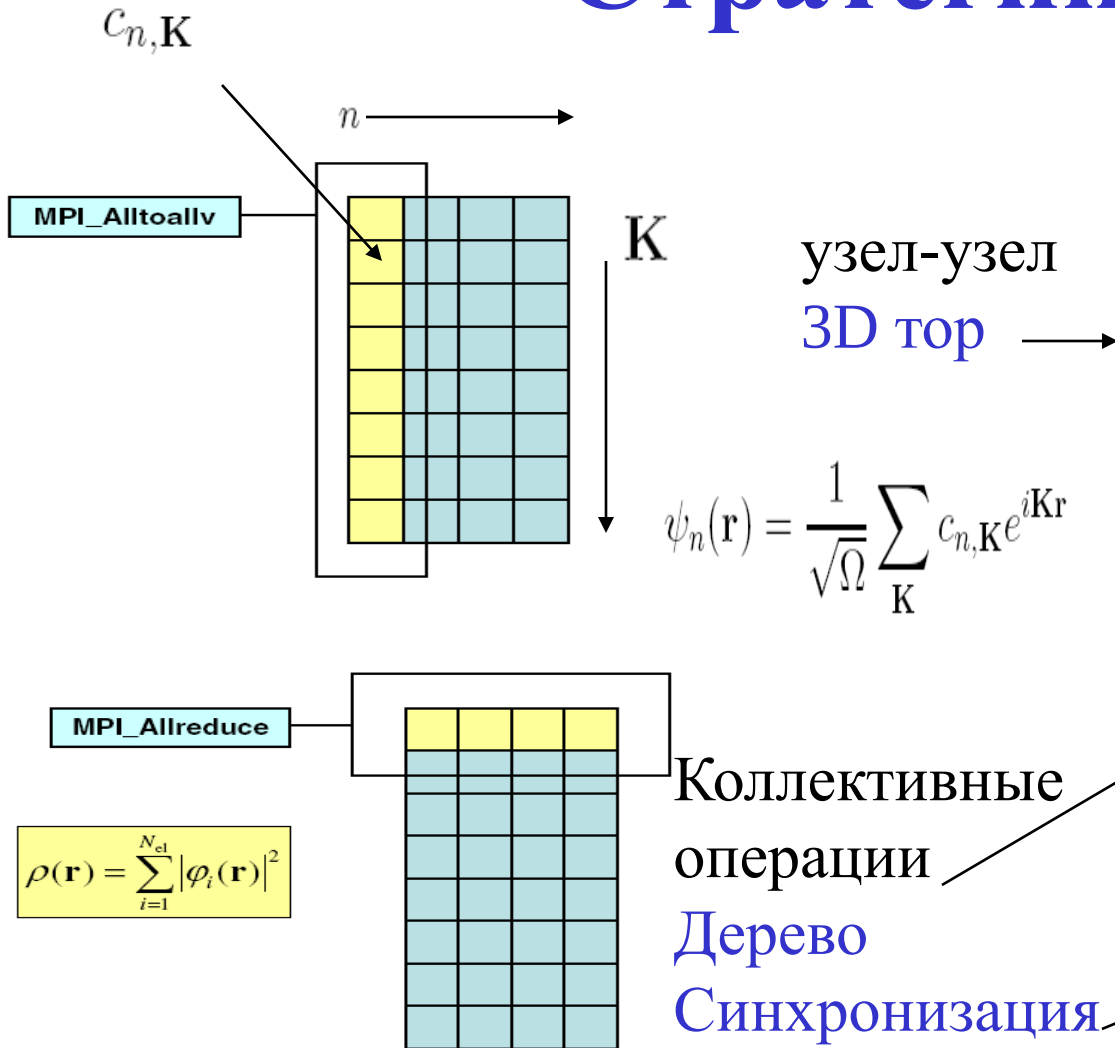
- **OpenMP** распараллеливание с общей памятью **на узле**. Длинные циклы.
- **Taskgroups** – группы процессоров. Процессоры организованы как двумерная сетка. Иерархическая параллелизация. Схема требует в два раза меньше коммуникаций, чем обычная схема. Процессорные группы могут быть оптимально распределены по тору.
- **Методы копирования**. Интегралы по траекториям в молекулярной динамике. Квантовые ядра. Многомерная метадинамика.

Блочное распределение коэффициентов Фурье волновых функций

- Блочное распределение коэффициентов Фурье волновых функций **C_nk** либо по орбитальному индексу “n” либо по индексу вектора обратной решетки “k”
- По “n” - вычисление Фурье- преобразования волновой функции на одном узле. Предел по размеру ячейки моделирования и размеру набора базисных функций из-за памяти на узел.
- По “k” - распределения базисного набора плоских волн. Условия для поддержки баланса загрузки. Усложнение выполнения 3d FFT. Коммуникации каждый с каждым.
- Прямая **декомпозиция в реальном пространстве** между процессорами (декомпозиция по слоям). Число процессоров не должно превышать наибольшего размера сетки (100-200) для того чтобы обеспечить баланс загрузки.
- Qbox опирается на библиотеки BLAS и ScaLAPACK
- CPMD – смешанная схема распараллеливания

Стратегии

МРІ
КОММУНИКАЦИИ
В ВЕР

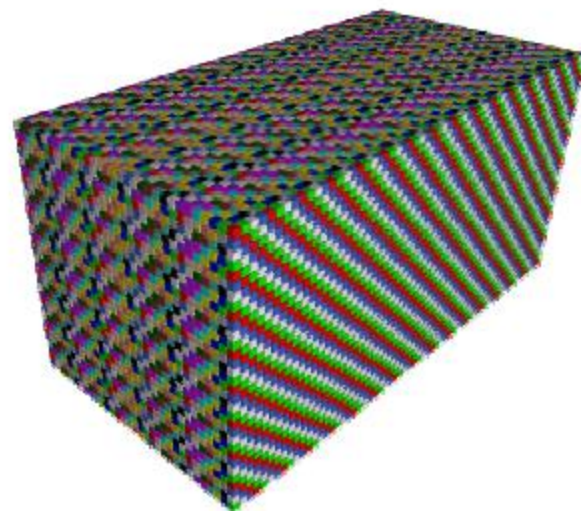
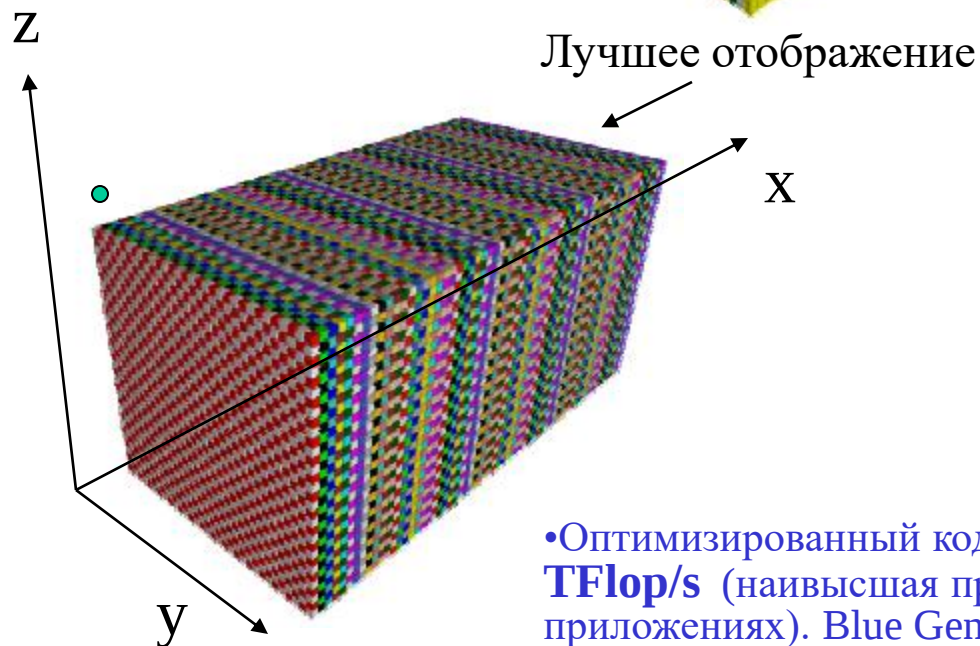
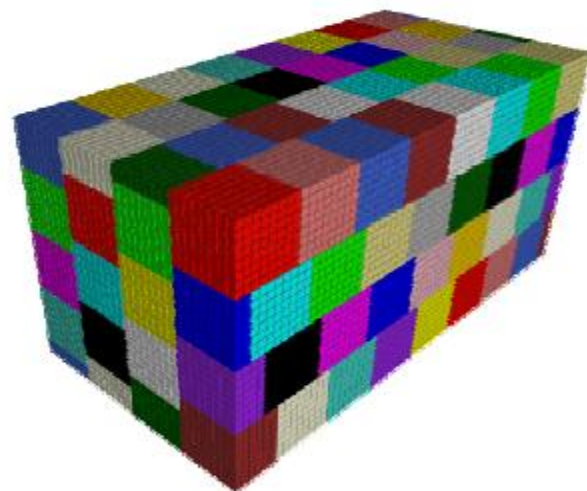
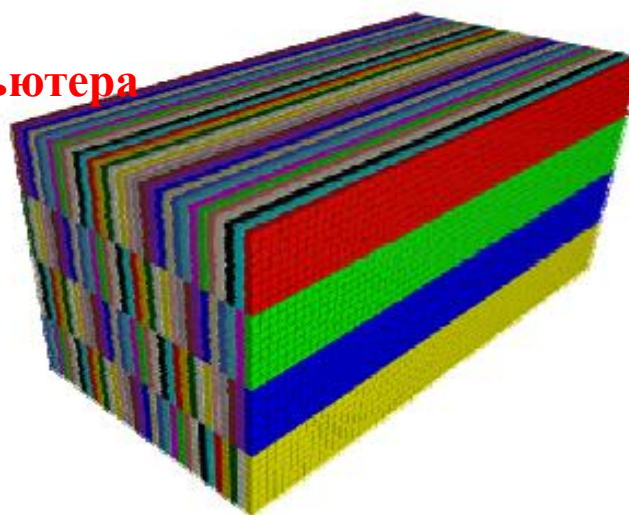


Блочное распределение коэффициентов Фурье волновых функций. Пространственная декомпозиция.

Отражение задач на архитектуру суперкомпьютера

- Для $P=65536$ (BGL) хороший выбор отображения дает 60% ускорения

- Оптимизация отображения - $N!$ отображений.



- Оптимизированный код Qbox показал производительность **207 TFlop/s** (наивысшая производительность на научных приложениях). Blue Gene/L

Производительность кода квантовой молекулярной

динамики на Blue Gene/P

- Вычисления на супер-ЭВМ IBM Blue Gene/P. Компьютер на факультете ВМК МГУ имеет **2048 процессорных узлов**, каждый узел - 4 ядра, пик - **27,8 терафлопс**.
- 82 атома –20 часов 512 узлов, 130 атомов –97 часов 512 узлов, 178 атомов –94 часа 1024 узла.
- **100 атомов масштабируются до 2048 узлов с 70% эффективности и длительностью производственного цикла ~ 600ps/неделя (МГУ)**

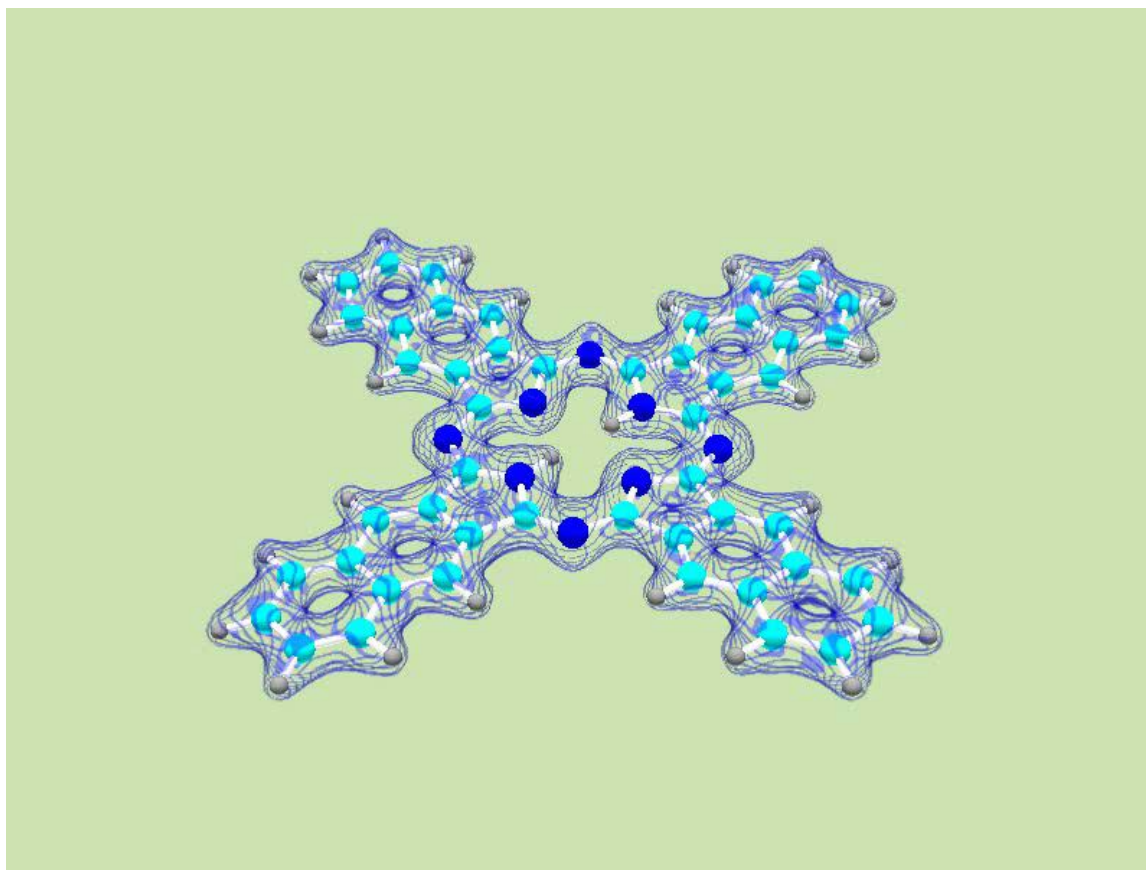
Масштабирование на петафлопные системы.

- Наибольшая система в настоящий момент **20000** атомов на 16 стойках.
- CRMD масштабируется до **128000** процессоров.
- Петафлопные компьютеры ~ **1000,000 !** процессоров. Необходима переделка кодов.
- Использование **многоядерных узлов** (тенденция в увеличении числа ядер)
- Производительность **ввода-вывода** – узкое место.
- Производительность параллельной файловой системы. Алгоритмы **сжатия данных**.
- Развитие **линейно-масштабируемых методов $O(N)$**
- **Надежность** компьютеров, сохранение работы программы при отказе узлов.

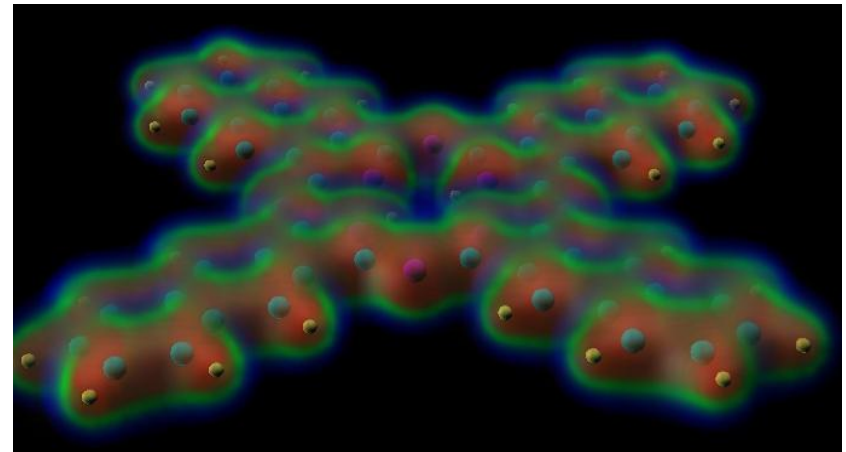
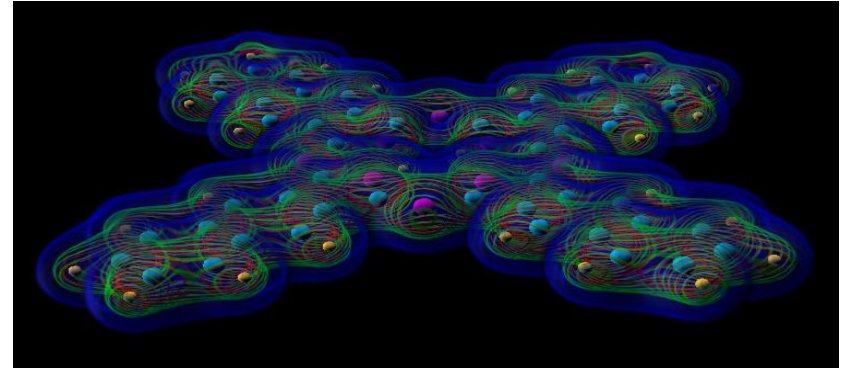
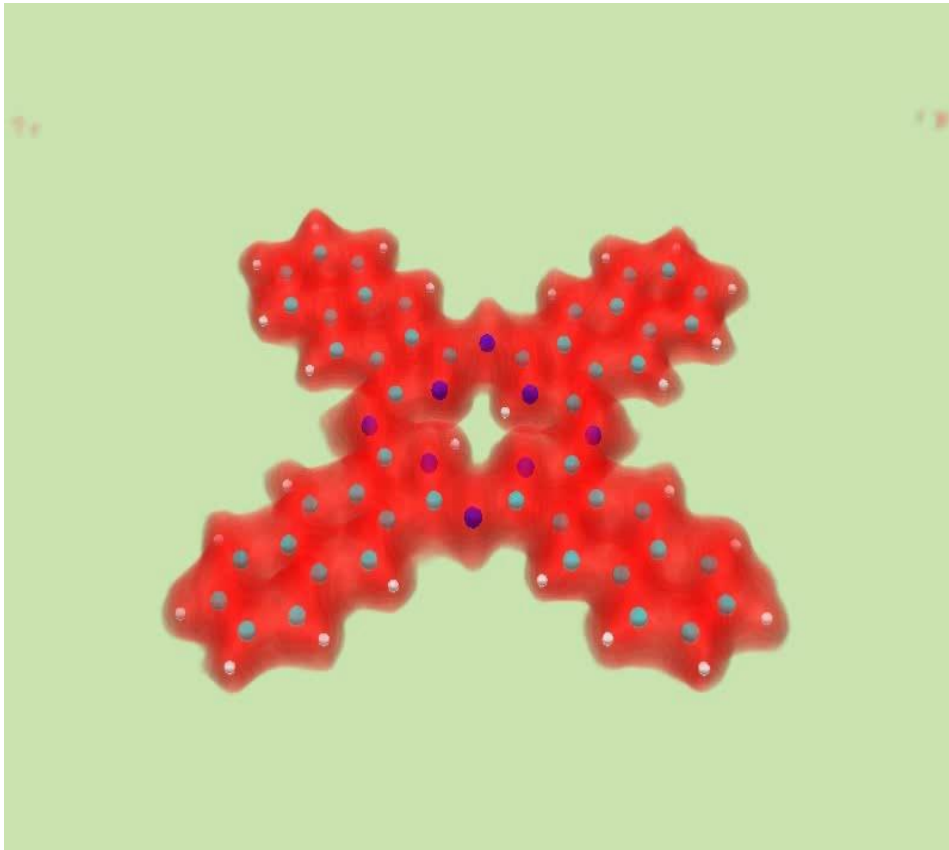
Создана параллельная система визуализации результатов научных исследований в реальном времени на Blue Gene/P МГУ.

Система предоставляет возможность выполнения визуализации в процессе работы параллельных прикладных программ.

Система поддерживает работу с данными большого объема и обеспечивает построение высококачественных изображений.



Параллельная визуализация в реальном времени на BG/P (МГУ)



Результаты

1. **Нанoeлектроника.** Представлено суперкомпьютерное моделирование из первых принципов молекулярного переключения на основе реакции изомеризации. Предложена **многомасштабная модель молекулярного переключения.** Сравнение с экспериментом.
2. **Эффективность суперкомпьютинга.** Создана параллельная **система визуализации** результатов научных исследований на BGP. Система поддерживает работу с данными большого объема и обеспечивает построение высококачественных изображений.
3. Разработан метод и **инструментальная система для настройки эффективности** параллельных программ, построенных на основе параллельных математических библиотек.