

**НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ**

Механико-математический факультет

Кафедра
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Зав. кафедрой, профессор

Глинский Борис Михайлович

Тел. (383)3306279, Факс (383)3308783,

E-mail: gbm@sccc.ru,

ИСТОРИЯ, ЗАДАЧИ

Кафедра была организована в 1987 году.

Первым заведующим кафедрой был член-корреспондент РАН В.Е. Котов.

Основной задачей кафедры является подготовка высокопрофессиональных специалистов широкого профиля по математическому моделированию, программному обеспечению, архитектуре ЭВМ и кластерных вычислительных систем, сетевым информационным системам.

Обучение навыкам практической работы на компьютерах, в том числе и на многоядерных кластерных суперЭВМ.

Специализация: **Математика и компьютерные науки,**
Прикладная математика и информатика

ОСНОВНЫЕ НАУЧНЫЕ НАПРАВЛЕНИЯ КАФЕДРЫ

1. Теория параллельных систем и процессов
(профессор Вирбицкайте И.Б.)
2. Суперкомпьютерные системы, кластерные суперЭВМ
(профессор Глинский Б.М.)
3. Моделирование, оптимизация и проектирование телекоммуникационных систем
(профессор Родионов А.С.)
4. Методы и технологии параллельных вычислений
(профессор Вшивков В.А., к.ф.-м.н.Куликов И.М., к.ф.-м.н. Снытников А.В.)
5. Технология проектирования больших программных систем и баз данных (доцент Васючкова Т.С.)
6. Параллельные системы обработки изображений и нейрокомпьютерные системы
(доцент Тарков М.С.)

СПИСОК СПЕЦКУРСОВ

Архитектура суперкомпьютерных системы (п/год, 1 сем),
д.т.н., профессор Глинский Б.М.

Теория параллельных процессов (год, 2 сем.),
д.ф.-м.н., профессор Вирбицкайте И.Б.

Технология программирования (год, 2 сем), к.ф.-м.н.,
доцент Васючкова Т.С.

Менеджмент и экономика программирования (п/год, 2 сем) ,
к.ф.-м.н., доцент Васючкова Т.С.

Моделирование и оптимизация телекоммуникационных систем (год), д.т.н. профессор Родионов А.С.

Нейрокомпьютерные системы (год), к.т.н., доцент Тарков М.С.

Распределенные алгоритмы (п/год), к.т.н., доцент Тарков М.С.

Квантовые компьютеры (п/год), к.т.н., доцент Тарков М.С.

Параллельные вычислительные методы (п/год),
д.ф.-м.н., профессор Вшивков В.А., к.ф.-м.н. Снытников А.В.

ОБЯЗАТЕЛЬНЫЕ СПЕЦСЕМИНАРЫ

- ◆ **Методы разработки больших программных систем (год)**
к.ф.- м.н. Васючкова Т.С.
- ◆ **Архитектура, системное и прикладное ПО кластерных суперЭВМ (год)**
д.т.н. Глинский Б.М.
- ◆ **Моделирование систем информатики (год)**
д.т.н. Родионов А.С.

СПЕЦКУРС, СЕМИНАР

Спецкурс: Суперкомпьютерные системы

Изучается архитектура, классы современных суперкомпьютеров. Основы построения и применения вычислительных кластеров на серверных платформах. Кластеры ЦКП ССКЦ.

Семинар: Архитектура, системное и прикладное программное обеспечение кластерных суперЭВМ

Обсуждаются технические, системные и прикладные аспекты современных кластерных систем для высокопроизводительных вычислений с участием представителей ведущих фирм: Intel, IBM, HP, Sun, Dell и др.

СОСТАВ ОСНОВНЫХ ТЕХНИЧЕСКИХ И ПРОГРАММНЫХ СРЕДСТВ ЦКП ССКЦ

КЛАСТЕР НКС-30Т+GPU гибридной архитектуры



576 процессоров (2688 ядер)

Intel Xeon E5450/E5540/X5670;

80 процессоров CPU (X5670) (480 ядер);

120 процессоров GPU - Tesla M 2090 (61440 ядер).

Общая пиковая производительность 2-х кластеров **115 Тфлопс**

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

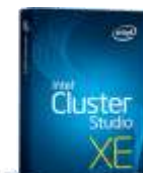
Общесистемное:

RedHat 5.4 - Операционная система, PBSPro 11.1 - Очередь заданий.

Средства разработки ПО (компиляторы, средства отладки, оптимизаторы программ):

Intel Parallel Studio XE 2013

Intel Cluster Studio XE 2013



СЕРВЕР с общей памятью (hp DL980 G7)



8 процессоров в (80 ядер) Intel E7-4870; ОП - 1 Тбайт

СИСТЕМЫ ХРАНЕНИЯ ДАННЫХ (СХД)



Кластерная файловая система IBRIX для НКС-30Т 32 Тбайта. **В сентябре 2014 г. добавлено 47 Тбайт.**

MPI 4.1 Intel, Intel TraceAnalyzer&Collector, Intel C++ & Fortran

NVIDIA CUDA 6.5, Portland Group PGI Accelerator 14.9.

Пакеты прикладных программ:

ANSYS CFD 14.5, Gaussian 09 GROMACS 4.6

Пакеты, разработанные в ИВМиМГ СО РАН:

PARMONC
AGNES
KPYLOV

Гибридный кластер НКС-30Т + GPU



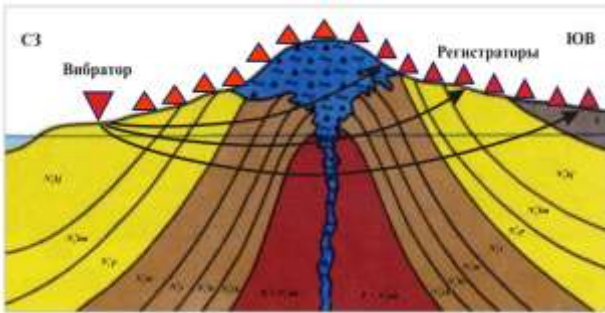
Архитектура ЦКП ССКЦ поддерживает две современных парадигмы параллельных вычислений – **MPI** для систем с распределенной памятью (МРР-кластеров) и **OpenMP** для систем с общей памятью.

Смешанная схема вычислений (**MPI+OpenMP**) позволяет запуск на каждый вычислительный узел кластера по одному **MPI**-процессу, который запускает внутри каждого вычислительного модуля несколько потоков с помощью **OpenMP**.

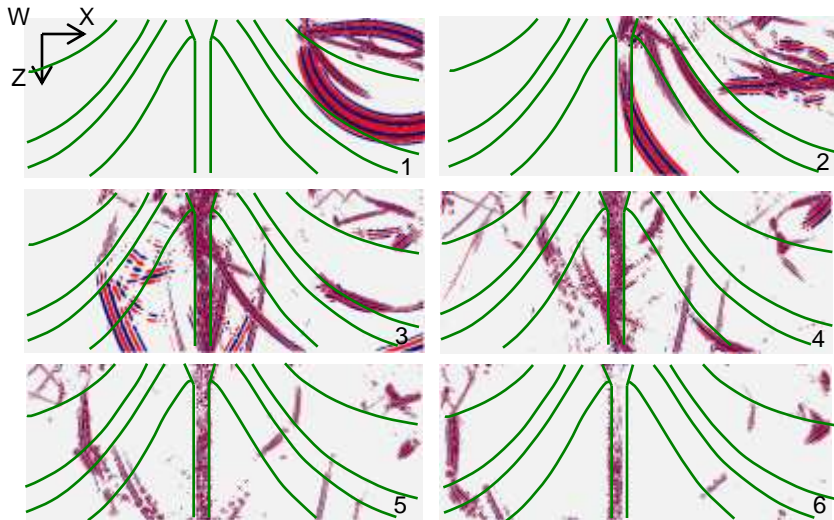
Для гибридной архитектуры: суперкомпьютер состоит из набора соединенных между собой узлов, для обмена данными используется **MPI**; каждый узел состоит из двух CPU и трёх GPU; на каждом узле запускается 1 процесс **MPI**, управляющий вычислениями (процесс выполняется на CPU); из **MPI** процесса запускаются потоки (threads) **OpenMP**, каждый из которых управляет работой одного GPU. Другой вариант: запускаются три **MPI** процесса на узел, каждый управляет закрепленным за ним GPU.

МОДЕЛИРОВАНИЕ 3D СЕЙСМИЧЕСКОГО ПОЛЯ ДЛЯ ВУЛКАНИЧЕСКИХ СТРУКТУР

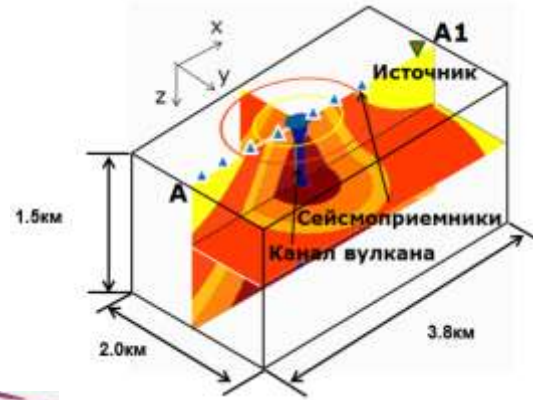
Геофизический натуральный эксперимент



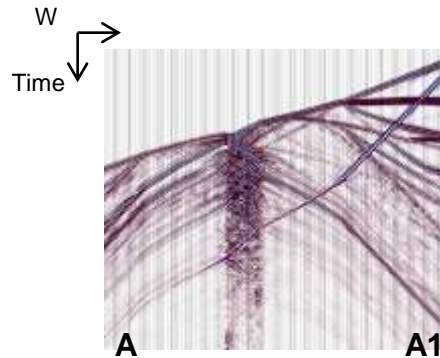
Синтетические снимки волнового поля



Геофизическая модель



Синтетическая сейсмограмма



Количество узлов 3D сетки:

X - 2096

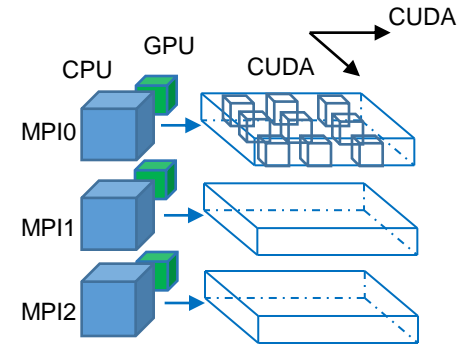
Y - 1103

Z - 828

$\sim 2.0 \cdot 10^9$

Более 20 3D массивов с данными

Параллельные вычисления

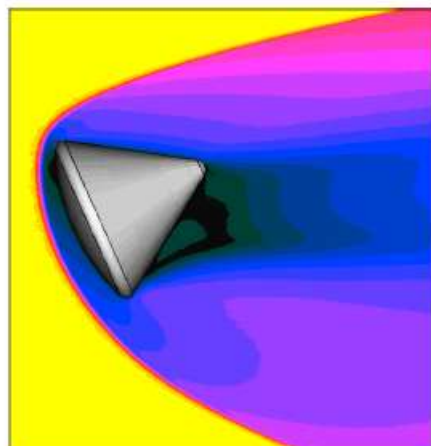


Кластер НКС-30Т+GPU ЦКП ССКЦ

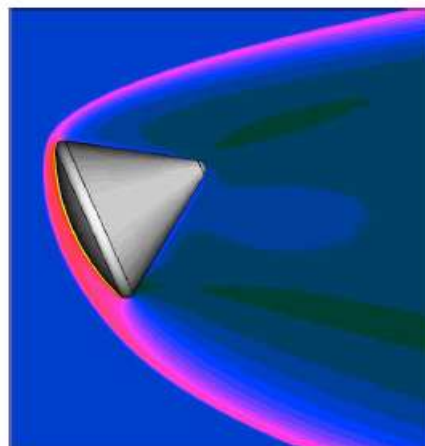
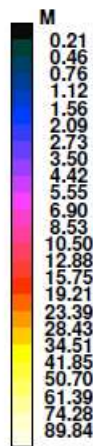
Спускаемый аппарат Аполлон

$H=85$ км; $V=9600$ м/с; $\alpha = -25^\circ$; смесь N_2, O_2

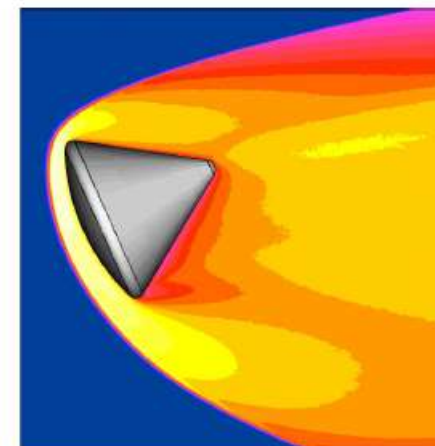
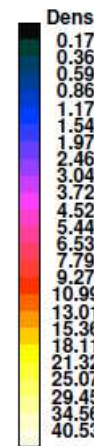
Кластер НКС-30Т, Сибирский суперкомпьютерный центр



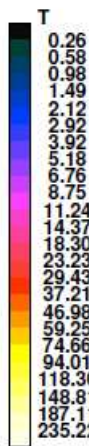
Число Маха



ρ/ρ_∞



T/T_∞



GPU	Частиц [млн]	Ячеек [млн]	Время счета	E_p [%]	S_p
24	705/794	62.5/103	9h00m	96.8	23.2
48	705/794	62.5/103	4h33m	96.3	46.2
48	1410/1590	62.5/215	9h39m	95.8	46



Синтэ-2 состоит из 16 тысяч узлов, каждый из которых включает в себя 2 процессора [Intel Xeon E5-2692](#) на архитектуре [Ivy Bridge](#) с 12 ядрами каждый (частота 2,2 ГГц) и 3 специализированных сопроцессора Intel [Xeon Phi 31S1P](#) (на архитектуре [Intel MIC](#), по 57 ядер, частота 1,1 ГГц, пассивное охлаждение). На каждом узле установлено 64 ГБ (16 модулей) оперативной памяти типа [DDR3 ECC](#) и дополнительно по 8 ГБ [GDDR5](#) в каждом узле (всего 88 ГБ). В общей сложности, количество вычислительных ядер достигает 3,12 миллиона (384 тысячи Ivy Bridge и 2736 тыс. Xeon Phi), производительность 33.863; 302 (Pflops)

ПРИМЕРЫ ПРИМЕНЕНИЯ СУПЕРКОМПЬЮТЕРОВ В ПРОМЫШЛЕННОСТИ

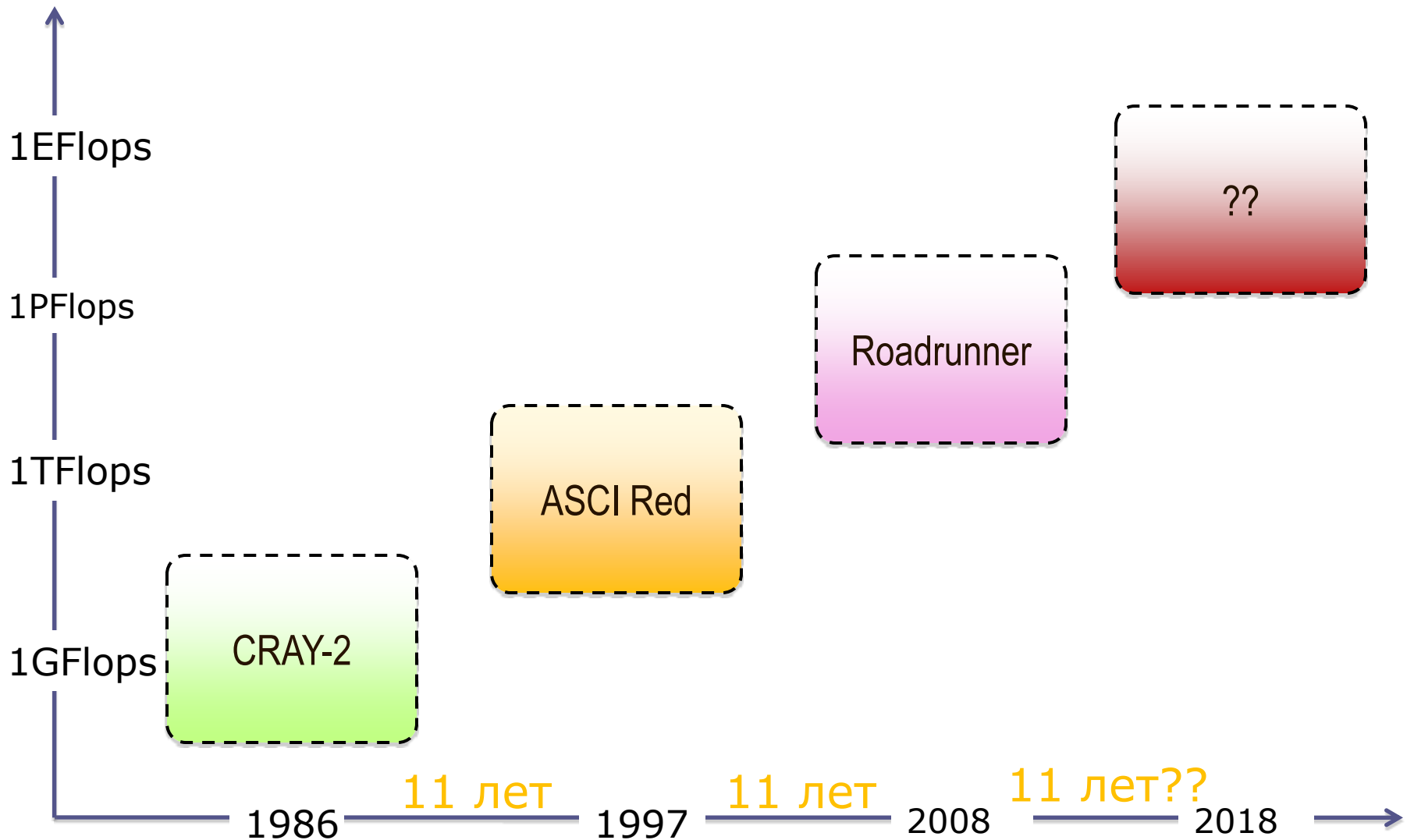
За рубежом

- Boeing 787 полностью сконструирован на компьютере. В 7 раз сокращено количество реальных испытаний. Разработка сокращена на год, экономия \$2миллиарда.
- Toyota CAMRY спроектирована на основе результатов численного моделирования, без проведения CRASH-тестов. Срок разработки сокращен с 30 до 8 месяцев.

В России

- НПО «Сатурн» проектирует на суперкомпьютерах авиационные двигатели.
- ЮрГУ используют вычислительный кластер для раскройки одежды, конструировании бронежилетов, в машиностроительной промышленности

ЭКЗАФЛОПСНЫЙ РУБЕЖ: КОГДА?



ЧТО ДАЛЬШЕ?

Прогноз Джека Донгарра

- 2018 г, производительность 1 эксафлопс (10^{18})
- Гибридная архитектура, интеграция CPU, GP-GPU на одном чипе
- Количество ядер 10-100 млн.

Проект IESP- International Exascale Software Project
(Д. Донгарра, П. Бекман).

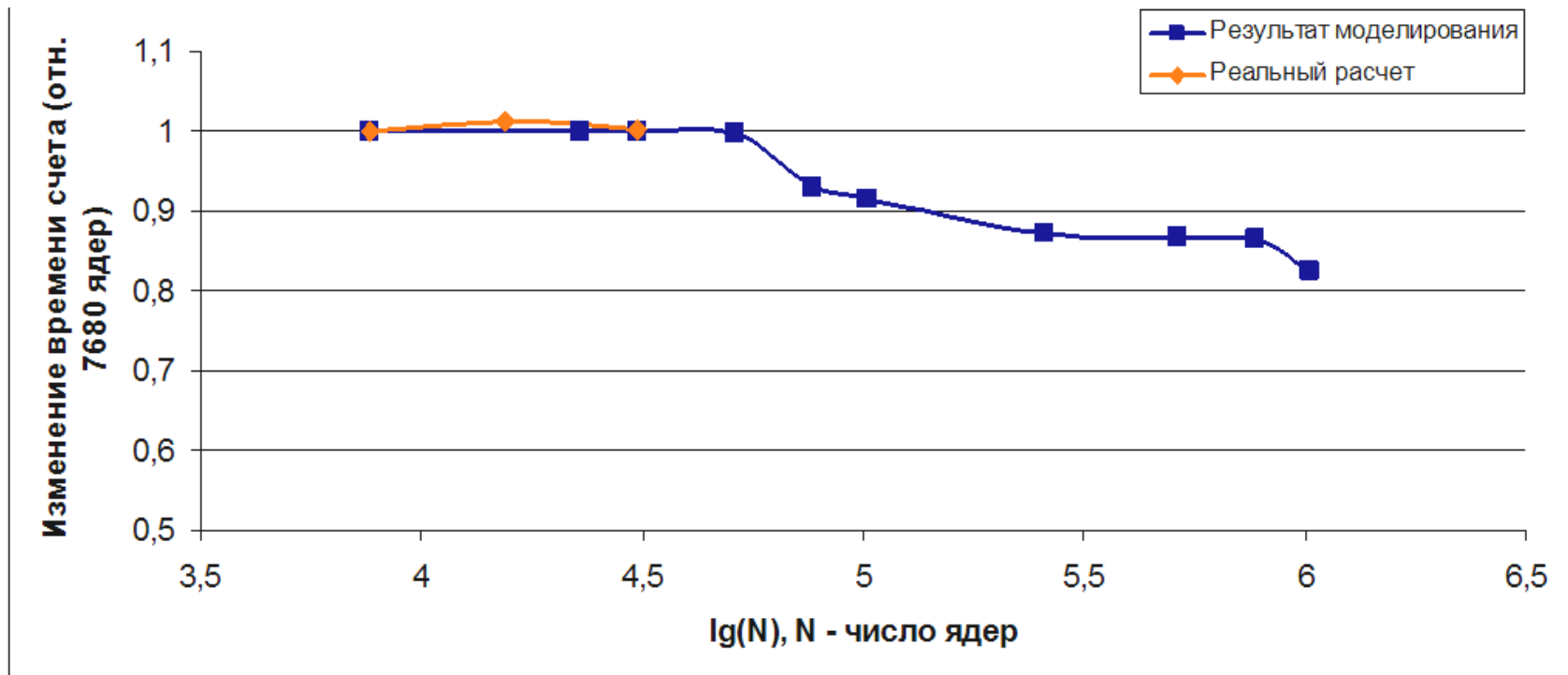
Проблемы

Переработка ОС, компиляторов, библиотек, сред программирования, методов и языков параллельных вычислений и др.

ПРОБЛЕМЫ МАСШТАБИРОВАНИЯ АЛГОРИТМОВ

- Исследование свойств масштабируемости параллельных алгоритмов является важной задачей при оценке эффективности их реализации, как для настоящих, так и будущих суперкомпьютеров пета- и эксафлопсного уровня.
- Данная проблема выходит за уровень технологических задач и требует научно-исследовательского подхода к ее решению.
- Вычислительные алгоритмы, как правило, являются более консервативными по сравнению с развитием средств вычислительной техники.
- Оценить поведение алгоритмов можно путем реализации на имитационной модели, отображающей их поведение на миллионах вычислительных ядер.
- Имитационная модель позволяет выявить узкие места в алгоритмах, понять, как нужно модифицировать алгоритм, какие параметры необходимо настраивать при его масштабировании на большое количество ядер при заданной архитектуре вычислительной системы.

РЕЗУЛЬТАТЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ (гибридный кластер)



Реальный расчет показан на начальном участке кривой (до 30720 ядер).
Модельный расчет (до 1 024 000 ядер).
При увеличении количества вычислительных узлов (от 7680 до 1 024 000 ядер) с пропорциональным увеличением размера 3D модели время счета увеличивается на 17,5%.

НАПРАВЛЕНИЕ ИССЛЕДОВАНИЙ

(темы дипломных работ)

- Исследование поведения алгоритмов решения прямых и обратных задач геофизики при их отображении методами имитационного моделирования на экзафлопсные суперкомпьютеры
- Разработка средств представления перспективных архитектур суперкомпьютеров в имитационных моделях.
- Выбор оптимальной архитектуры экзафлопсной ЭВМ для решения различных классов задач

SEQUOIA суперкомпьютер Ливерморской национальной лаборатории



Количество ядер	1 572 864
Общий объем системной памяти	1.5 Pб
Общая пиковая производительность	20 Пфлопс
Рабочая производительность (Linpack)	16.32 Пфлопс
Электропотребление системы в Квт	7890

«К» суперкомпьютер, RIKEN Advanced Институт вычислительных наук (AICS), Япония



Количество ядер	705 024
Общая пиковая производительность	11. 280 Пфлопс
Рабочая производительность (Linpack)	10 .510 Пфлопс
Электропотребление системы в Квт	12 659.9

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ СЕТОЧНОГО МЕТОДА С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ AGNES (AGent NEtwork Simulator)

- Область исследования режется вдоль одной оси, и полученные области загружаются на вычислители. У каждого вычислителя есть пересечение по данным максимум с 2-ми вычислителями.
- Функциональные агенты: Grid — узел-вычислитель, имитирующий расчет сеточных методов на одном вычислителе.
- Масштабируемость – увеличивая количество вычислителей, увеличиваем размерность задачи, загружая на новые вычислители новые области такого же размера как и на предыдущих вычислителях. Ожидаем сохранение общего времени счета, при фиксированном количестве шагов алгоритма на одном вычислителе.