

СУПЕРКОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ

Б.М. Глинский

Институт Вычислительной Математики и Математической Геофизики СО РАН

ОСНОВНОЙ КОЛЛЕКТИВ

ИВМиМГ СО РАН:

- Б.М. Глинский
- И.М. Куликов
- А.Ф. Сапетина
- И.Г. Черных
- Д.В. Винс
- М.А. Марченко
- А.С. Родионов
- А.В. Снытников
- П.А. Титов
- Г.Ф. Жерняк

ИСИ СО РАН:

- Ю.А. Загорулько
- Г.Б. Загорулько

Работа выполнена при поддержке грантов РФФИ № 16-07-00434 и № 19-07-00085

Руководитель: Б.М. Глинский

СОДЕРЖАНИЕ

- АКТУАЛЬНОСТЬ РАБОТЫ
- ПРЕДЛАГАЕМЫЙ ИНТЕГРАЛЬНЫЙ ПОДХОД
- СО-ДИЗАЙН ПАРАЛЛЕЛЬНЫХ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ
- ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ
- ЭНЕРГОЭФФЕКТИВНОСТЬ АЛГОРИТМОВ
- ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА РЕШЕНИЯ ЗАДАЧ
- ЗАКЛЮЧЕНИЕ

АКТУАЛЬНОСТЬ РАБОТЫ

- Сложность применения современных и будущих экзафлопсных суперкомпьютеров связана с невероятной степенью параллелизма, это сотни миллионов и миллиарды одновременно работающих вычислительных ядер.
- Уровень развития супервычислений для машин этого класса связан как с разработкой архитектур самих суперкомпьютеров, так и с совершенствованием их алгоритмического и программного обеспечения.
- Для каждой предметной области необходимым становится выбор метода решения задачи и соответствующей многоядерной архитектуры наиболее эффективно решающего задачу за приемлемое время с заданной точностью.
- Актуальным становится разработка новых вычислительных технологий основанных на со-дизайне, имитационном моделировании, оценке энергоэффективности алгоритмов и интеллектуальной поддержке решения задачи.

СУПЕРКОМПЬЮТЕРЫ



Стойки с оборудованием Summit соединяют 300 км оптоволоконного кабеля, а 15 000 литров циркулируют охлаждая их. Вычислительная мощность Summit составляет 122 петафлопс, потенциально максимальная – 200 петафлопс. Ориентирован на задачи ИИ.

Компании HPE и AMD в 2023 году планируют запустить суперкомпьютер El Capitan. Производительностью 2 экзафлопса. Ориентирован на задачи с интенсивной работой с данными, такими как моделирование, симуляция, аналитика и работа с ИИ.

ИНТЕГРАЛЬНЫЙ ПОДХОД

Суперкомпьютерная технология предложенная авторами для разработки алгоритмического и программного обеспечения содержит три связанных этапа:

- 1) со-дизайн, под которым мы понимаем адаптацию постановки задачи, математического метода, вычислительного алгоритма под параллельную архитектуру суперкомпьютера на всех этапах решения задачи;
- 2) оценка масштабируемости вычислительного алгоритма для суперкомпьютеров на основе имитационного моделирования;
- 3) оценка энергоэффективности алгоритмов при различных реализациях на заданной архитектуре суперЭВМ.

КО-ДИЗАЙН МЕТОДОВ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО-СЛОЖНЫХ ЗАДАЧ

Понятие ко-дизайна в зарубежной литературе часто применяется при разработке архитектуры компьютеров (*Ang J. et al. High Performance Computing: From Grids and Clouds to Exascale, chapter Exascale Computing and the Role of Co-design. IOS PressInc, 2011*).

Этот путь является заведомо тупиковым, так как не учитывает развитие программного обеспечения для подобных архитектур (*Barret R. et al. On the Role of Co-design in High Performance Computing. Transition of HPC Towards Exascale Computing. IOS Press, 2013*).

В последнее время под ко-дизайном понимается совместная разработка архитектуры и программного обеспечения (Hardware/Software co-design).

Для реализации такой концепции созданы специальные рабочие группы в лаборатории Лос-Аламоса (<http://codesign.lanl.gov/>) и Ливерморской лаборатории (<https://codesign.llnl.gov/>).

С 2011 года проводятся международные конференции по различным сторонам концепции ко-дизайна (<http://www.ncic.ac.cn/codesign/>).

СО-ДИЗАЙН МЕТОДОВ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО-СЛОЖНЫХ ЗАДАЧ

Авторы доклада предлагают расширить этот термин применительно к разработке алгоритмического и программного обеспечения суперкомпьютеров.

В данном случае **основой понятия со-дизайна является анализ эффективности параллельной реализации на всех этапах решения задачи под заданную архитектуру суперкомпьютера:**

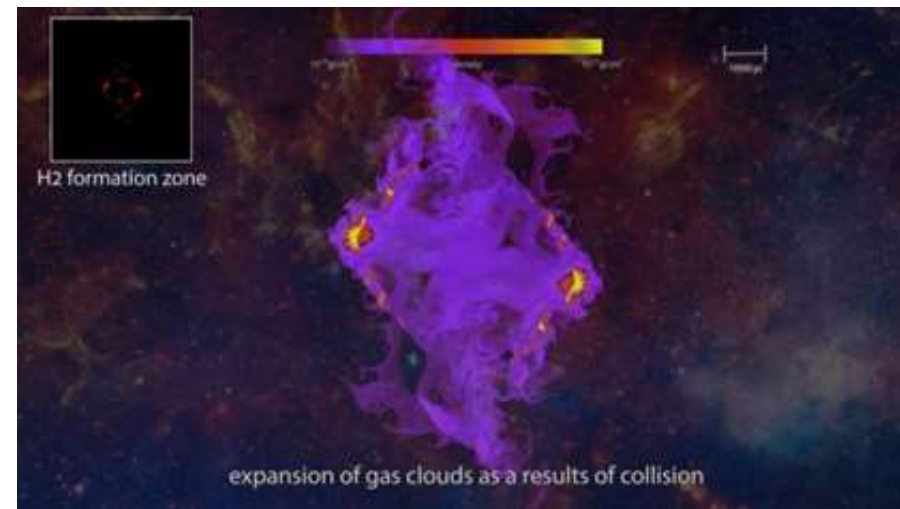
- 1) формулировка «физической» постановки задачи;
- 2) математическая формулировка «физической» задачи;
- 3) создание численного метода для решения математической постановки;
- 4) выбор структур данных и параллельного алгоритма;
- 5) учёт особенностей архитектуры суперЭВМ;
- 6) использование средств разработки.

Данный подход применялся при решении задач астрофизики, геофизики, физике плазмы.

КОНЦЕПЦИЯ СО-ДИЗАЙНА

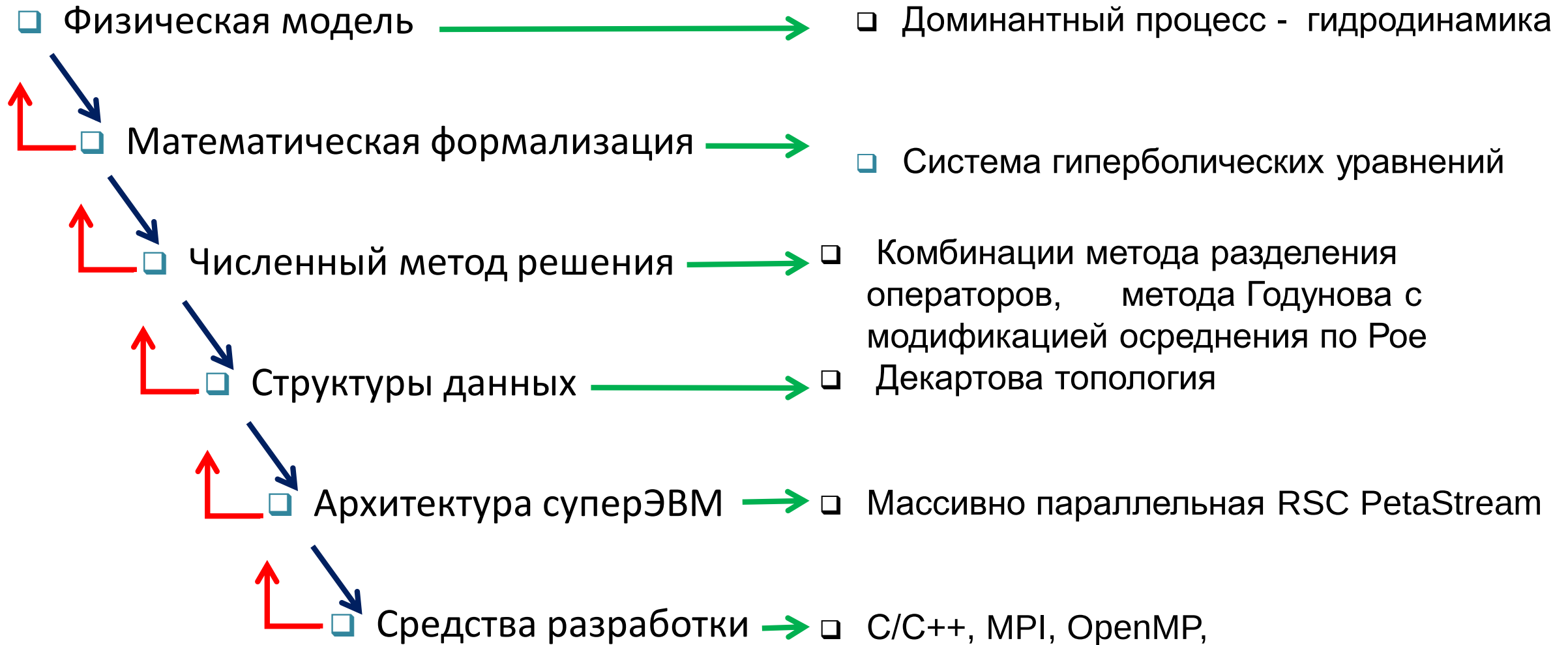


МНОГОКОМПОНЕНТНАЯ ГИДРОДИНАМИЧЕСКАЯ МОДЕЛЬ СТОЛКНОВЕНИЯ ГАЛАКТИК



Химодинамика столкновения галактик: начальная стадия (слева), разлет газовых облаков после столкновения и зона образования молекулярного водорода (справа). Одна галактика пролетает через другую с образованием двух газовых облаков и зоны образования молекулярного водорода после удара.

РЕАЛИЗАЦИЯ СО-ДИЗАЙНА ДЛЯ АСТРОФИЗИЧЕСКОГО МОДЕЛИРОВАНИЯ



ЧИСЛЕННЫЙ МЕТОД РЕШЕНИЯ АСТРОФИЗИЧЕСКОЙ ЗАДАЧИ

- Operator splitting подход (разделение на эйлеров и лагранжев этапы)
- Метод Годунова для каждого этапа
- Специальная модификация схемы Рунге для осреднения величин
- Кусочно-параболический метод на локальном шаблоне
- Быстрое преобразование Фурье
- Использование переопределенной системы уравнений (гарантия неубывания энтропии)

В частности, код AstroPhi основан на описанных методах. Для его реализации используется архитектура ускорителей Intel Xeon Phi. Для моделирования используется архитектура RSC PetaStream с 8 узлами с 64 ускорителем Intel Xeon Phi 7120D.

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ АЛГОРИТМОВ

Исследование масштабируемости параллельных алгоритмов проводится с применением мультиагентного распределённого моделирования систем с дискретными событиями AGNES (AGent based NEtwork Simultor), разработанной в ИВМиМГ СО РАН.

Основу системы представляют три группы агентов: системные, отвечающие за функционирование модели в конкретной операционно-вычислительной среде; управляющие, отвечающие за отказоустойчивое имитационное моделирование, включая управление синхронизацией событий и сбора статистики; проблемно-ориентированные, имитирующие работы компонентов моделируемых программ и суперЭВМ.

МУЛЬТИАГЕНТНАЯ СИСТЕМА AGNES, ОСНОВАННАЯ НА ПЛАТФОРМЕ JADE

Используемой платформой JADE разделение функционала агента на набор поведений позволяет легко распараллелить исполнение независимых поведений.

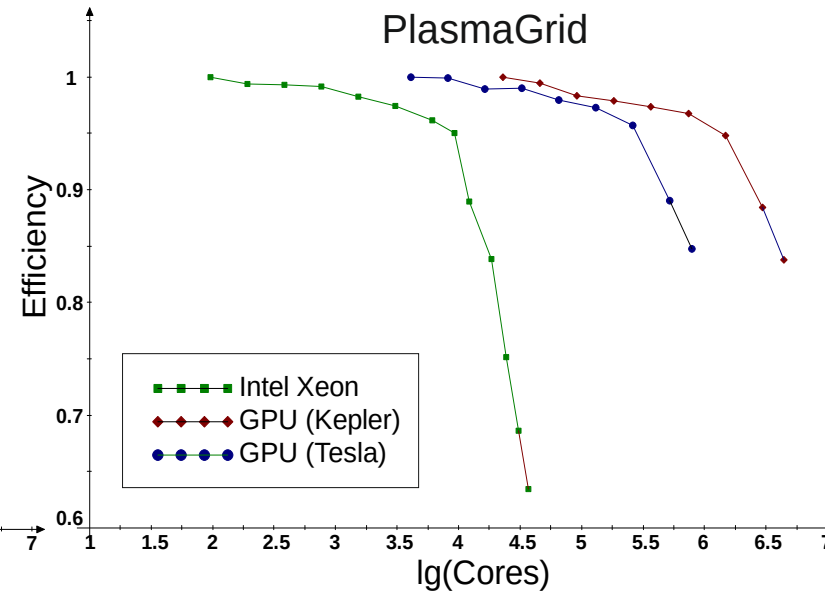
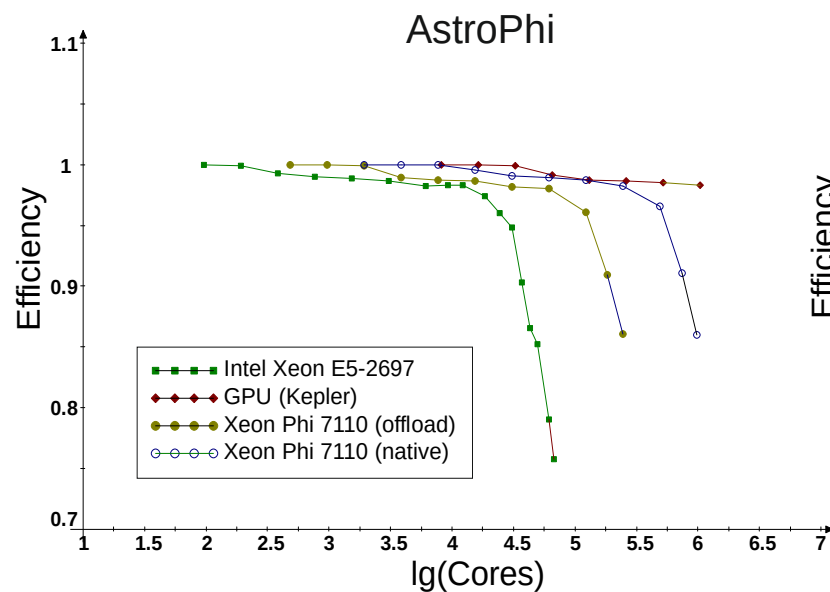
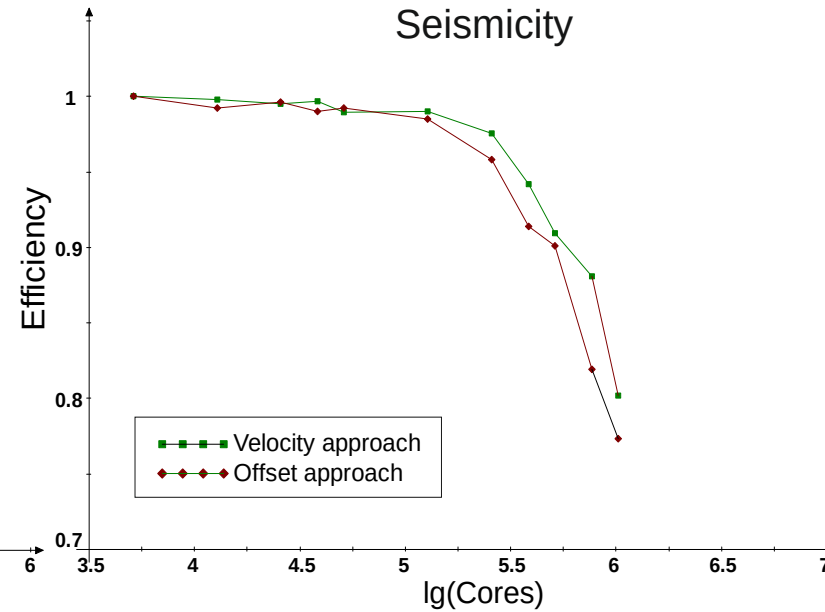
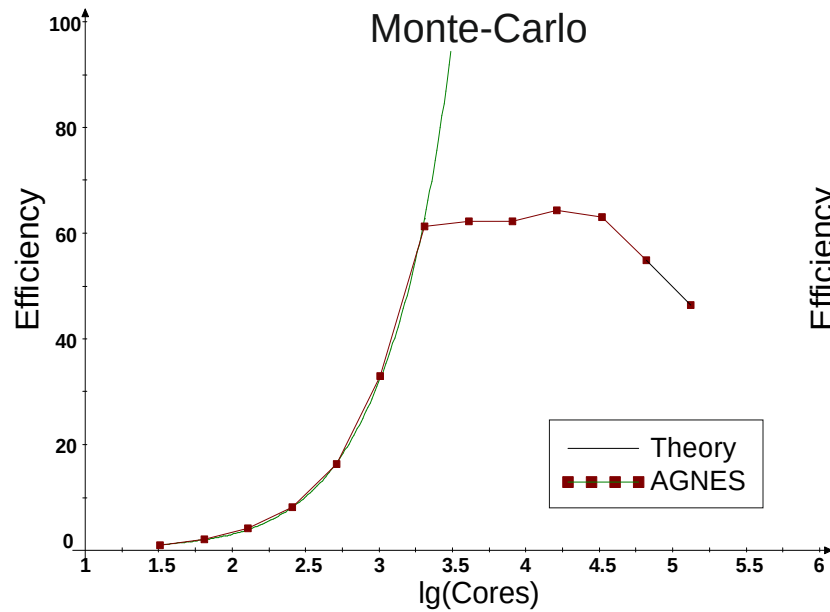
Платформа JADE также предусматривает разбиение агентов внутри платформы по контейнерам, которые могут находиться на различных хостах. Таким образом, появляется возможность производить распределенные вычисления и балансировать нагрузку.

Однако, при моделировании систем где «все агенты связаны со всеми» система будет работать все медленнее и медленнее.

Практика показывает, что при моделировании с AGNES на одном среднем вычислительном узле одновременный запуск более 1000-2000 простых взаимодействующих агентов вызывает заметное падение производительности.

В настоящее время в нашей лаборатории разрабатывается система имитационного моделирования на основе Акторов. Обмен короткими, асинхронными сообщениями в модели Акторов позволит получить более точные данные о масштабируемости алгоритма

РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ С ПОМОЩЬЮ СИСТЕМЫ AGNES



Для параллельных сеточных методов снижение эффективности при увеличении вычислительных ядер связано с лавинообразным ростом обменов сообщениями между ядрами, а для распределенного статистического моделирования тем, что узел-сборщик не успевает обрабатывать большое количество сообщений с промежуточными данными.

ЭНЕРГОЭФФЕКТИВНОСТЬ АЛГОРИТМОВ И ПРОГРАММ

- Параллельно развитию энергоэффективности аппаратных платформ, актуально исследование энергоэффективности алгоритмов и программ, которые сводятся к вычислительной эффективности, эффективности обработки данных, информированности о контексте вычислительной среды.
- С целью повышения энергоэффективности алгоритмов для задач сейсмики, астрофизики и физики плазмы каждый программный код оптимизировался по следующим трем направлениям: эффективность работы с процессором и оперативной памятью, эффективность работы с сетевыми устройствами и эффективность работы с периферическими системами.
- Для оптимизации эффективной работы с процессором и памятью использовался пакет Intel Parallel Studio XE 2016 Beta, включающий в себя Intel Vectorization Advisor, который анализирует код и выдает предложения для оптимизации программы с целью максимизации эффективности использования векторных инструкций. Использование Intel Vectorization Advisor позволило ускорить астрофизический код на 30%.
- Для задач сейсмики, астрофизики и физики плазмы были выбраны наиболее эффективные численные методы с точки зрения минимизации сетевых обменов.

ЭНЕРГОЭФФЕКТИВНОСТЬ

Профиль исполнения астрофизического кода на рабочем узле SMP-G7 (Intel E7-4870) ССКЦ

Name	TSelf	TSelf	TTotat	#Calls	TSelf /Call
▲ Group Application	154.396 s		157.082 s	4	38.5989 s
Process 0	38.4594 s		39.2725 s	1	38.4594 s
Process 1	38.9564 s		39.2733 s	1	38.9564 s
Process 2	38.9702 s		39.2756 s	1	38.9702 s
Process 3	38.0096 s		39.2609 s	1	38.0096 s
▷ Group MPI	2.68675 s		2.68675 s	384	6.99675e-3 s

Профиль исполнения астрофизического кода на рабочем узле G8-K40 (Intel Xeon E5-2650v2) ССКЦ

Name	TSelf	TSelf	TTotat	#Calls	TSelf /Call
▲ Group Application	44.1507e+3 s		45.0657e+3 s	16	2.75942e+3 s
Process 0	2.75706e+3 s		2.81676e+3 s	1	2.75706e+3 s
Process 1	2.7155e+3 s		2.81672e+3 s	1	2.7155e+3 s
Process 2	2.73732e+3 s		2.81674e+3 s	1	2.73732e+3 s
Process 3	2.73277e+3 s		2.81675e+3 s	1	2.73277e+3 s
Process 4	2.73852e+3 s		2.81672e+3 s	1	2.73852e+3 s
Process 5	2.70039e+3 s		2.81675e+3 s	1	2.70039e+3 s
Process 6	2.73803e+3 s		2.81674e+3 s	1	2.73803e+3 s
Process 7	2.75461e+3 s		2.81677e+3 s	1	2.75461e+3 s
Process 8	2.78787e+3 s		2.8166e+3 s	1	2.78787e+3 s
Process 9	2.78898e+3 s		2.81653e+3 s	1	2.78898e+3 s
Process 10	2.78098e+3 s		2.81635e+3 s	1	2.78098e+3 s
Process 11	2.78907e+3 s		2.81662e+3 s	1	2.78907e+3 s
Process 12	2.78067e+3 s		2.81623e+3 s	1	2.78067e+3 s
Process 13	2.79598e+3 s		2.8167e+3 s	1	2.79598e+3 s
Process 14	2.77751e+3 s		2.81614e+3 s	1	2.77751e+3 s
Process 15	2.77545e+3 s		2.81654e+3 s	1	2.77545e+3 s
▷ Group MPI	914.946 s		914.946 s	1776	515.172e-3 s

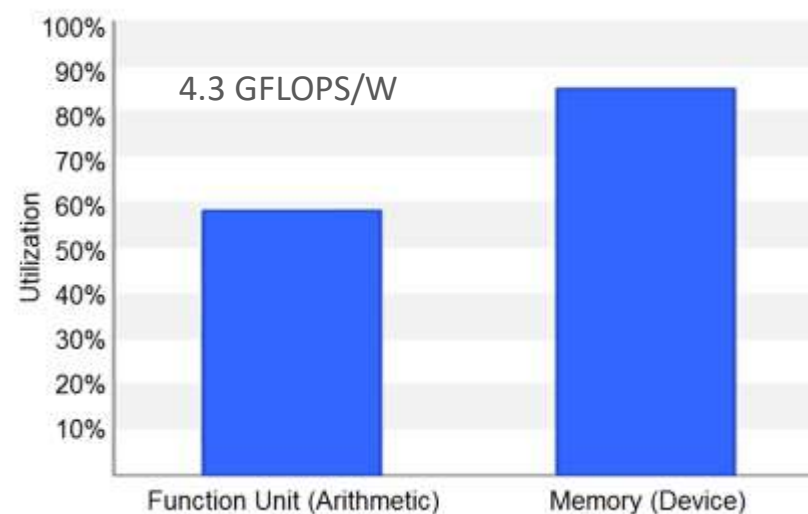
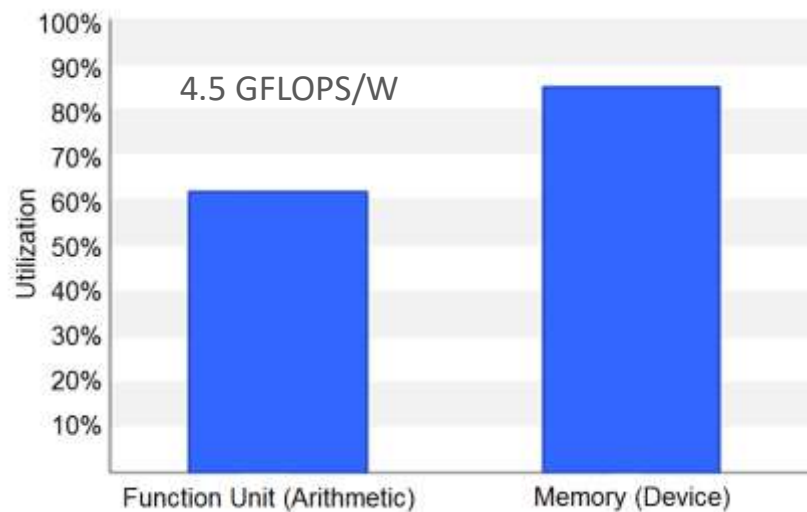
Оптимизация численных методов астрофизического кода и использование последних версий программных продуктов из пакета Intel Parallel Studio XE 2016 Beta дали отличный результат балансировки нагрузки на ядрах процессоров. Время выполнения MPI операций составляет не более 2-3 % от общего времени счета.

ОЦЕНКА ЭНЕРГОЭФФЕКТИВНОСТИ СЕЙСМИЧЕСКОГО КОДА

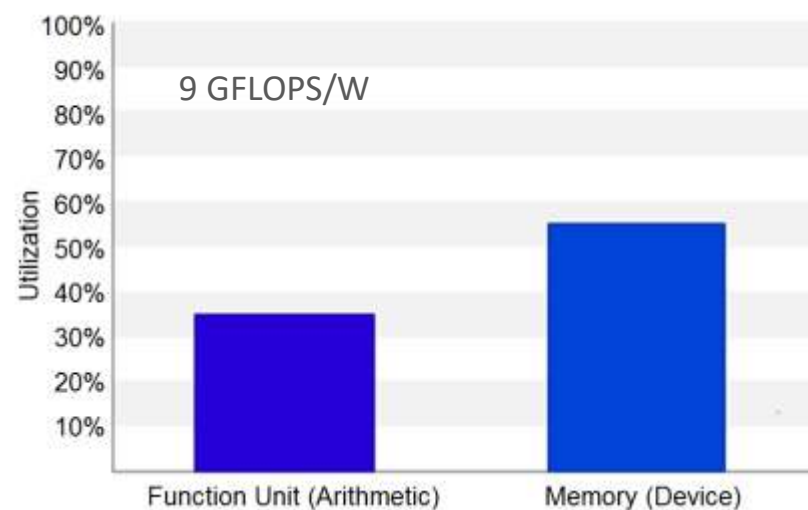
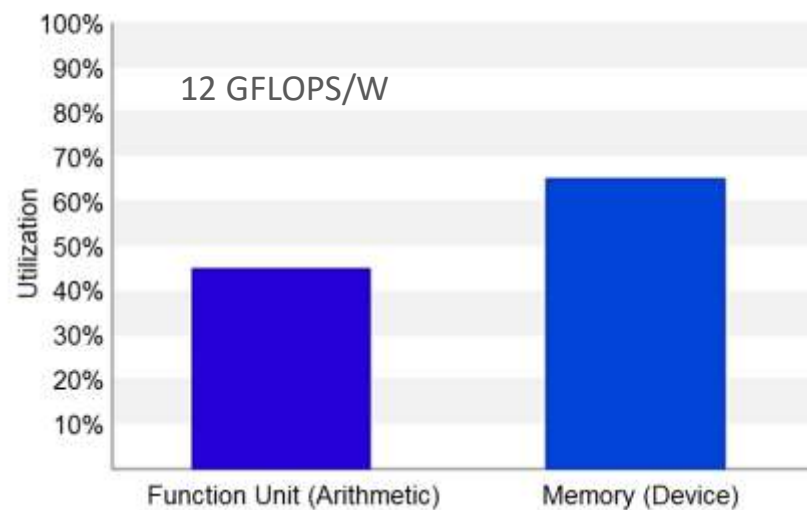
$\vec{\sigma}, \vec{u}$

$\vec{U}$

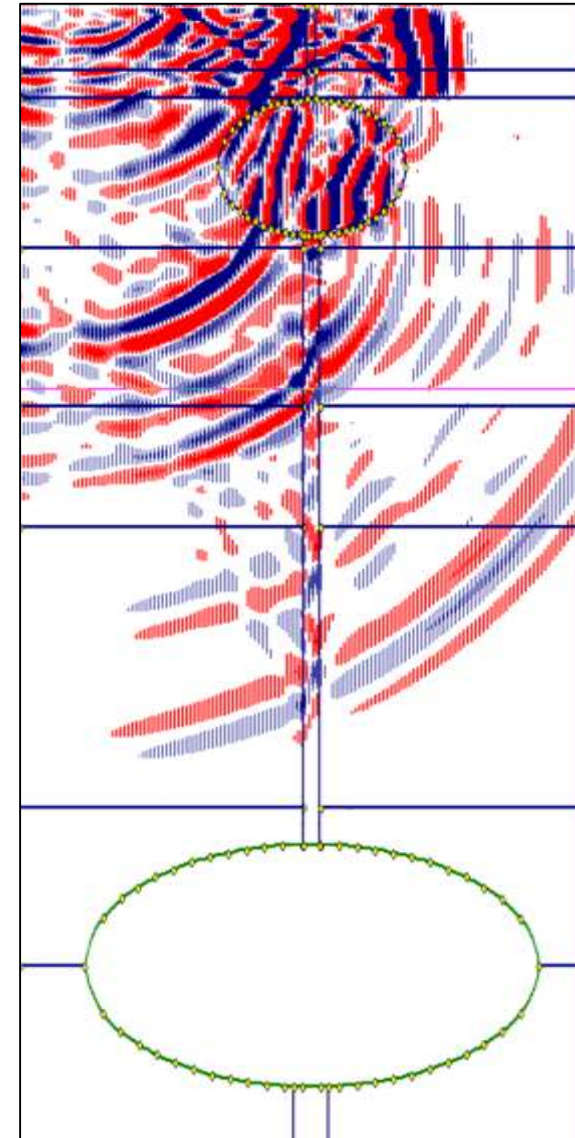
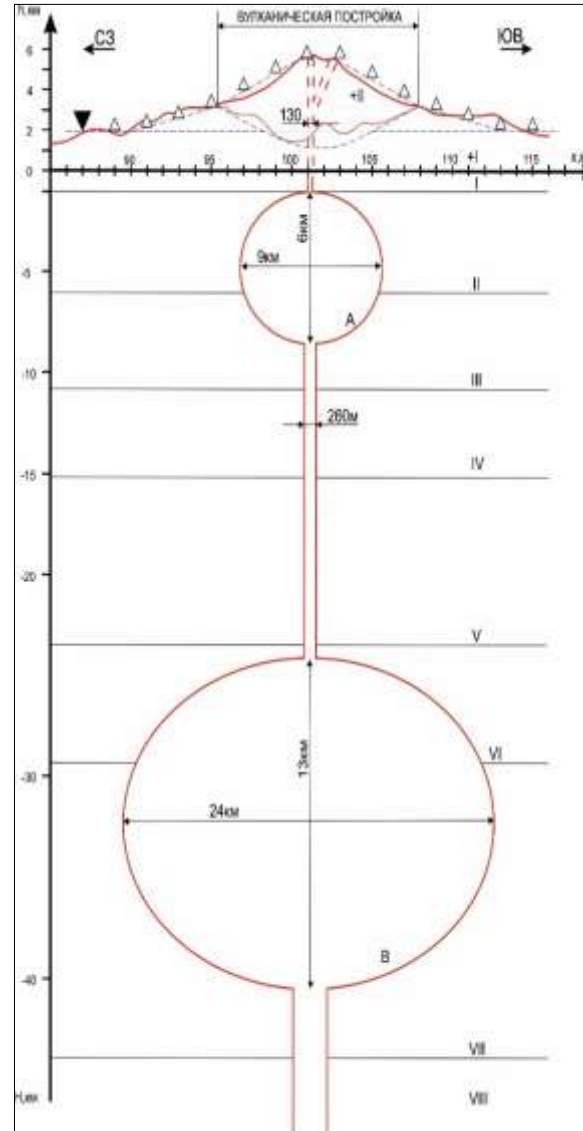
Tesla M2090



K40M

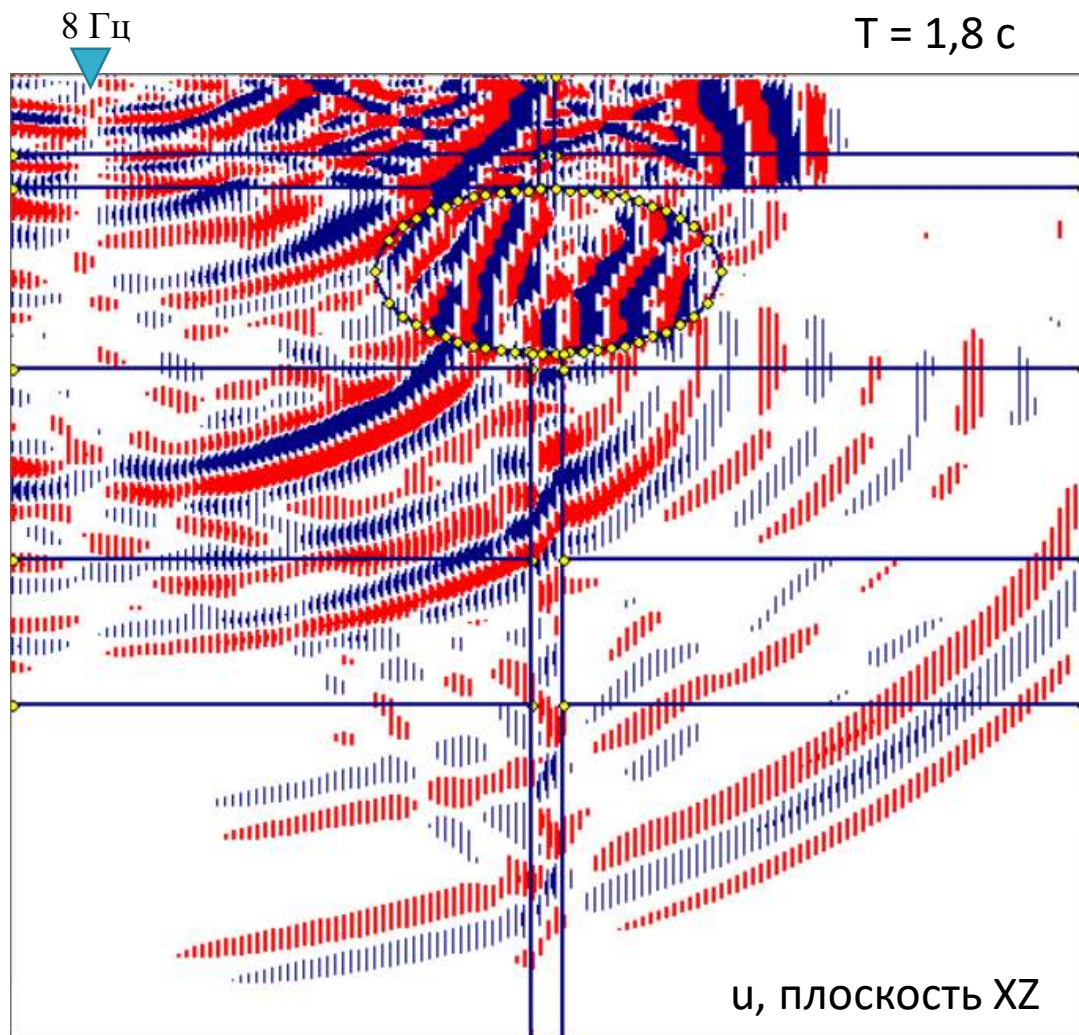


МОДЕЛИРОВАНИЕ РАСПРОСТРАНЕНИЯ СЕЙСМИЧЕСКИХ ВОЛН В СРЕДАХ, ХАРАКТЕРНЫХ ДЛЯ ВУЛКАНИЧЕСКИХ СТРУКТУР



ТЕСТОВАЯ МОДЕЛЬ

Интервал глубин, км	ρ , г/см ³	V_p , км/с	V_s , км/с
0 – 2	2,66	3,1	1,79
2 – 3	2,7	3,2	1,82
3 – 8	2,85	5,9	3,42
8 – 13	2,62	6,22	3,59
13 – 17	2,7	5,82	3,37
17 – 26	2,75	5,97	3,45
камеры и каналы	2,1	2,2	0



ПОСТАНОВКА ЗАДАЧИ В ТЕРМИНАХ СКОРОСТЕЙ СМЕЩЕНИЯ И НАПРЯЖЕНИЯ

$$\rho \frac{\partial \vec{u}}{\partial t} = [A] \vec{\sigma} + \vec{F}(t, x, y, z), \quad \frac{\partial \vec{\sigma}}{\partial t} = [B] \vec{u},$$

$$A = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} & 0 \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} & 0 & \frac{\partial}{\partial z} \\ 0 & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix}, \quad B = \begin{bmatrix} (\lambda + 2\mu) \frac{\partial}{\partial x} & \lambda \frac{\partial}{\partial y} & \lambda \frac{\partial}{\partial z} \\ \lambda \frac{\partial}{\partial x} & (\lambda + 2\mu) \frac{\partial}{\partial y} & \lambda \frac{\partial}{\partial z} \\ \lambda \frac{\partial}{\partial x} & \lambda \frac{\partial}{\partial y} & (\lambda + 2\mu) \frac{\partial}{\partial z} \\ \mu \frac{\partial}{\partial y} & \mu \frac{\partial}{\partial x} & 0 \\ \mu \frac{\partial}{\partial z} & 0 & \mu \frac{\partial}{\partial x} \\ 0 & \mu \frac{\partial}{\partial z} & \mu \frac{\partial}{\partial y} \end{bmatrix},$$

$$\sigma|_{t=0}=0, \quad u|_{t=0}=0, \quad \sigma_{xz}|_{z=0}=0, \sigma_{yz}|_{z=0}=0, \sigma_{zz}|_{z=0}=0$$

$\vec{u} = (u, v, w)^T$ – вектор скоростей смещений,

$\vec{\sigma} = (\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz})^T$ – тензор напряжений,

$\rho(x, y, z)$ – плотность, $\lambda(x, y, z), \mu(x, y, z)$ – параметры Ламе,

$\vec{F}(t, x, y, z)$ – массовая сила

ПОСТАНОВКА ЗАДАЧИ В ТЕРМИНАХ СМЕЩЕНИЙ

$$\rho \frac{\partial^2 \vec{U}}{\partial t^2} = [C] \vec{U} + \vec{F}(t, x, y, z)$$

$$C = \begin{bmatrix} (\lambda + 2\mu) \frac{\partial^2}{\partial x^2} + \mu \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) & (\lambda + \mu) \frac{\partial^2}{\partial x \partial y} & (\lambda + \mu) \frac{\partial^2}{\partial x \partial z} \\ (\lambda + \mu) \frac{\partial^2}{\partial y \partial x} & (\lambda + 2\mu) \frac{\partial^2}{\partial y^2} + \mu \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) & (\lambda + \mu) \frac{\partial^2}{\partial y \partial z} \\ (\lambda + \mu) \frac{\partial^2}{\partial z \partial x} & (\lambda + \mu) \frac{\partial^2}{\partial z \partial y} & (\lambda + 2\mu) \frac{\partial^2}{\partial z^2} + \mu \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \end{bmatrix}$$

$$U(x, y, z)|_{t=0} = 0, \quad \sigma_{xz}|_{z=0} = 0, \quad \sigma_{yz}|_{z=0} = 0, \quad \sigma_{zz}|_{z=0} = 0$$

$\vec{U} = (U, V, W)^T$ – тензор напряжений,

$\rho(x, y, z)$ – плотность, $\lambda(x, y, z), \mu(x, y, z)$ – параметры Ламе,

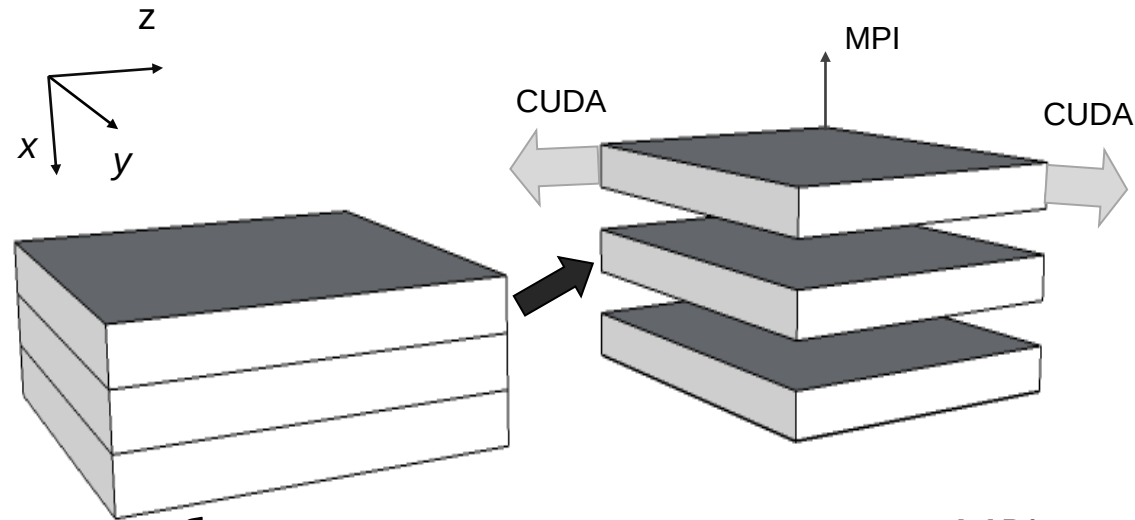
$\vec{F}(t, x, y, z)$ – массовая сила

МЕТОД РЕШЕНИЯ

- Конечно-разностный метод
- Явные конечно-разностные схемы на сдвинутых сетках*
- Второй порядок аппроксимации по времени и пространству
- Равномерные сетки
- Критерий устойчивости:
$$\tau \leq \left(Vp_{\max} \sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}} \right)^{-1},$$

где $Vp_{\max}$ – максимальная скорость распространения упругих волн

ПАРАЛЛЕЛЬНАЯ РЕАЛИЗАЦИЯ АЛГОРИТМОВ: ДЕКОМПОЗИЦИЯ И ОПТИМИЗАЦИЯ РАБОТЫ С ПАМЯТЬЮ



- Обмены между узлами: неблокированная передача данных MPI
- Обмены между графическими картами: асинхронные функции CUDA
- Оптимальное расположение всех используемых трехмерных массивов и соответствующее распределение работ между нитями.
- Сохранение в константную память графической карты всех основных констант, используемых на каждом шаге по времени.
- В памяти GPU хранятся не исходные сеточные коэффициенты, а их модификации. Реализация для **расчета скоростей смещения и напряжения** требует выделения памяти для хранения **26** трехмерных массивов, а реализация для **расчета смещений всего 14 массивов**.

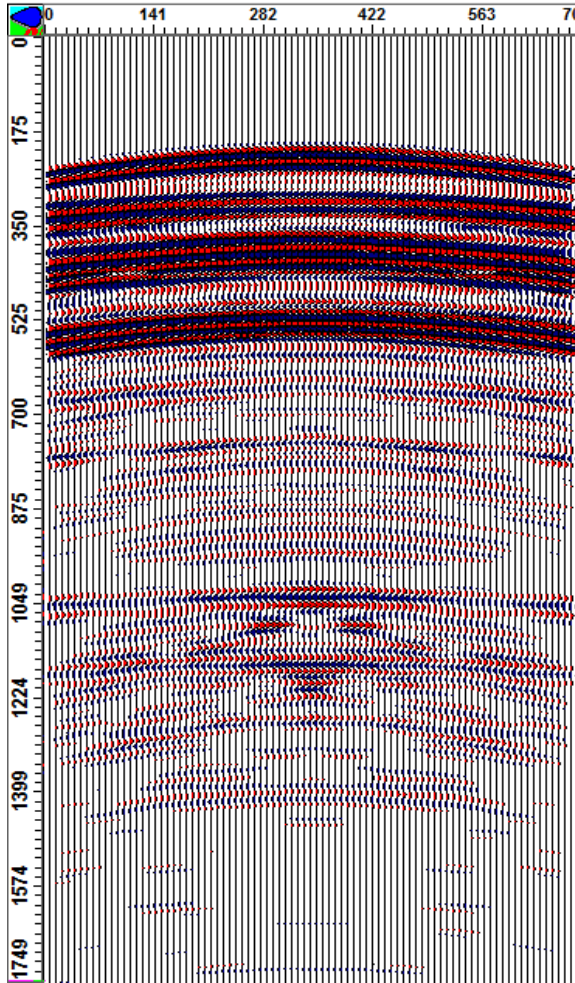
СРАВНЕНИЕ ВРЕМЕНИ РАСЧЕТОВ

	$\vec{\sigma} = (\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz})^T$ $\vec{u} = (u, v, w)^T$	$\vec{U} = (U, V, W)^T$
Кол-во уравнений	9	3
Кол-во 3D массивов в программе	26	14
Кол-во операций с плавающей точкой на один временной шаг	57	98
Время расчета на 15 узлах кластера НКС-30Т+GPU, с	183,1	174,8
Время расчета на 8 узлах кластера НКС-30Т+GPU, с		247,4

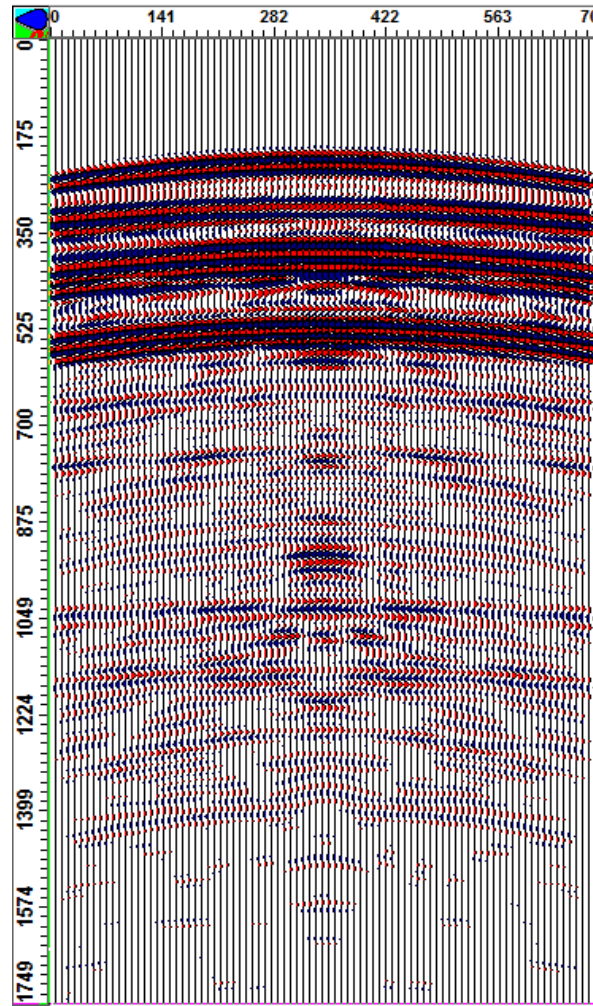
Кол-во узлов расчетной сетки	по оси X	1500
	по оси Y	700
	по оси Z	2100
	по времени	1000

СРАВНЕНИЕ 3D МОДЕЛИРОВАНИЯ СРЕД С ОТКРЫТЫМ И ЗАКУПОРЕННЫМ КАНАЛОМ ИЗВЕРЖЕНИЯ

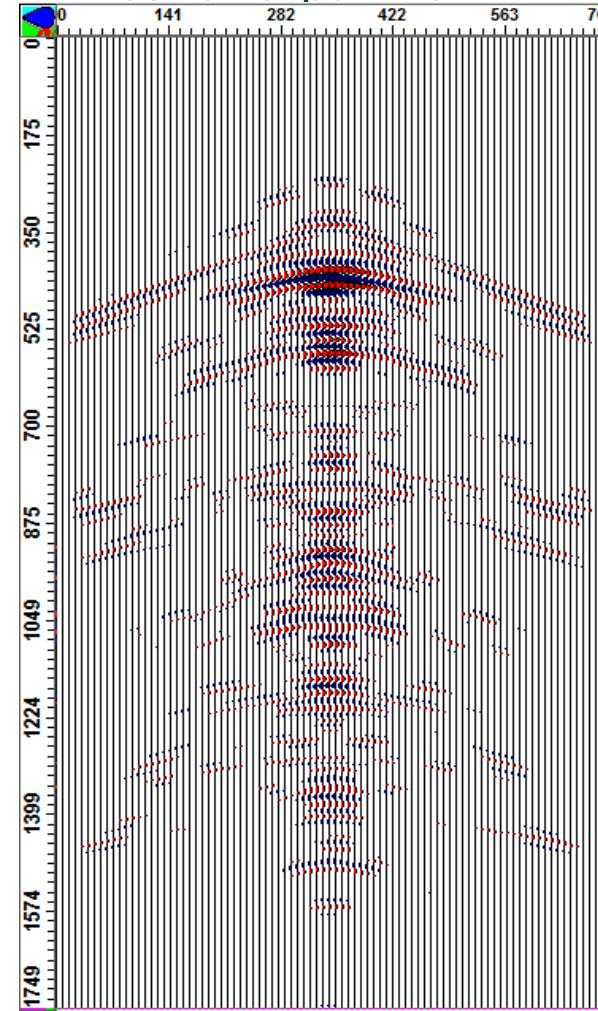
Равновесное состояние



Момент извержения



Разница между
сейсмограммами



ОБСУЖДЕНИЕ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ РЕШЕНИЯ СЕЙСМИЧЕСКОЙ ЗАДАЧИ

Для сравнения быстродействия обоих подходов в рамках концепции со-дизайна проведены расчеты одинаковых сред с размером сетки $1500*700*2100$ узлов по пространству и 1000 шагов по времени.

Взята сетка близкая к максимально входящей в память 45 графических карт (15 вычислительных узлов кластера НКС-30Т+GPU) при расчете скоростей смещения и напряжения.

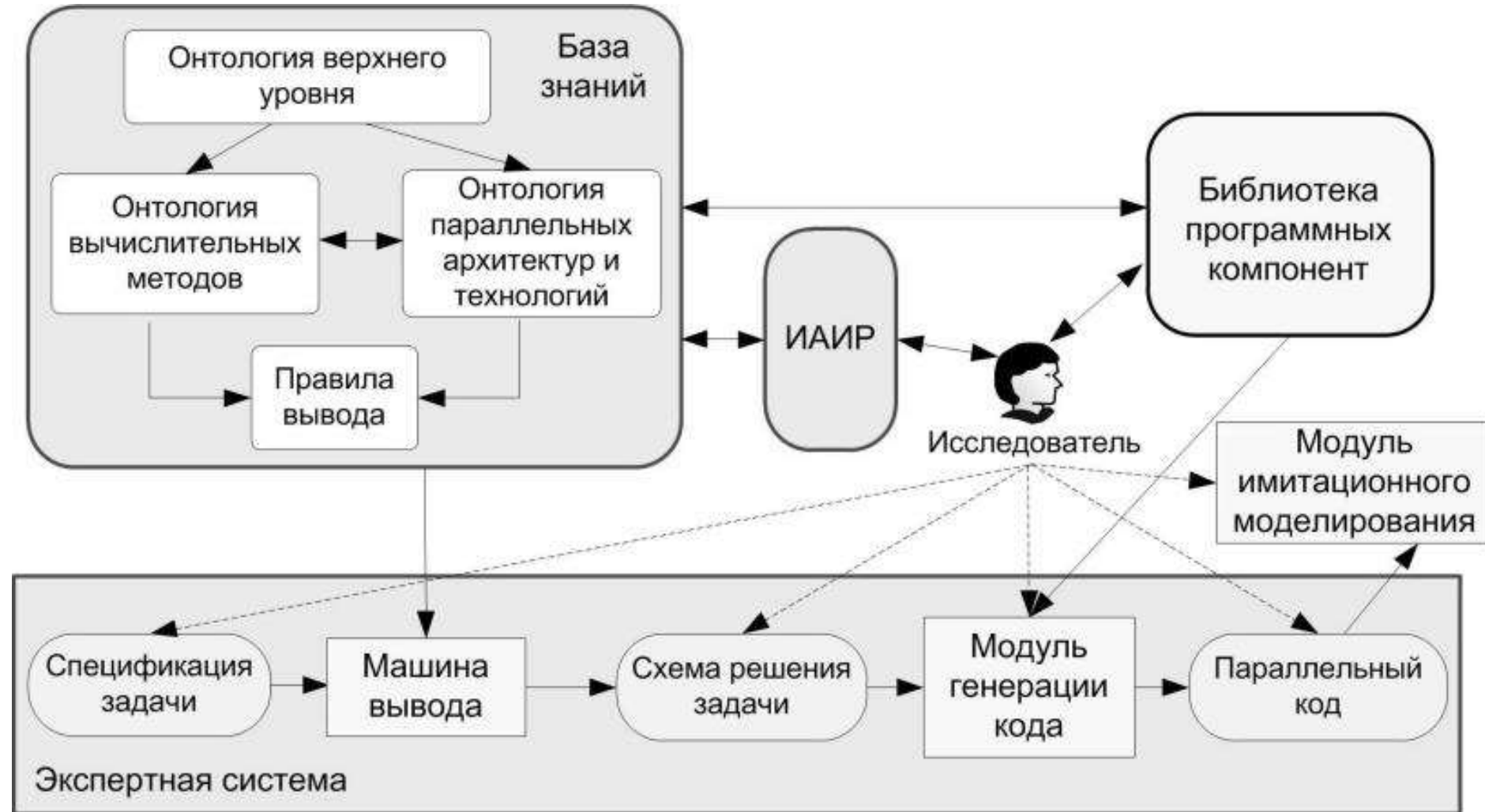
При расчете смещений на такой же по размеру сетке используется практически в два раза меньше оперативной памяти, и он может быть проведен минимум на 8 узлах вместо 15.

Проведенные эксперименты демонстрируют, что подход, основанный на расчете смещений, быстрее и при этом позволяет проводить расчеты для очень больших сеток, запрашивая меньшее количество свободных узлов. Это позволяет быстрее получать доступ к ним в условиях очереди на кластере, обеспечивая при этом приемлемое время расчета (несколько часов для реальной полномасштабной задачи).

ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ФИЗИЧЕСКИХ ЗАДАЧ

- Разработка программ для суперкомпьютеров существенно усложняется с наращиванием параллелизма и использованием гетерогенных архитектур с большой вычислительной мощностью.
- Актуальной задачей становится создание средств интеллектуальной поддержки решения вычислительно сложных задач как на современных, так и будущих суперкомпьютерах. Создание своего рода «моста» между пользователем и средствами решения задачи.
- За рубежом проводятся работы по созданию ресурсов, основанных на онтологических описаниях в различных областях науки (физика [Cvjetkovic, 2017], геология [Xiaogang, 2011], биология [Cook, 2013], астрофизика [Sarro, 2003; Louge, 2015] и др.). В настоящее время существуют три web-ресурса по астрофизике [ESPAS, 2019], по геномике [SemGen, 2019], по геологии и геофизике [Awesome geoscience semantics, 2019].
- Авторы данной работы предлагают использовать онтологический подход к решению вычислительно сложных физических задач.

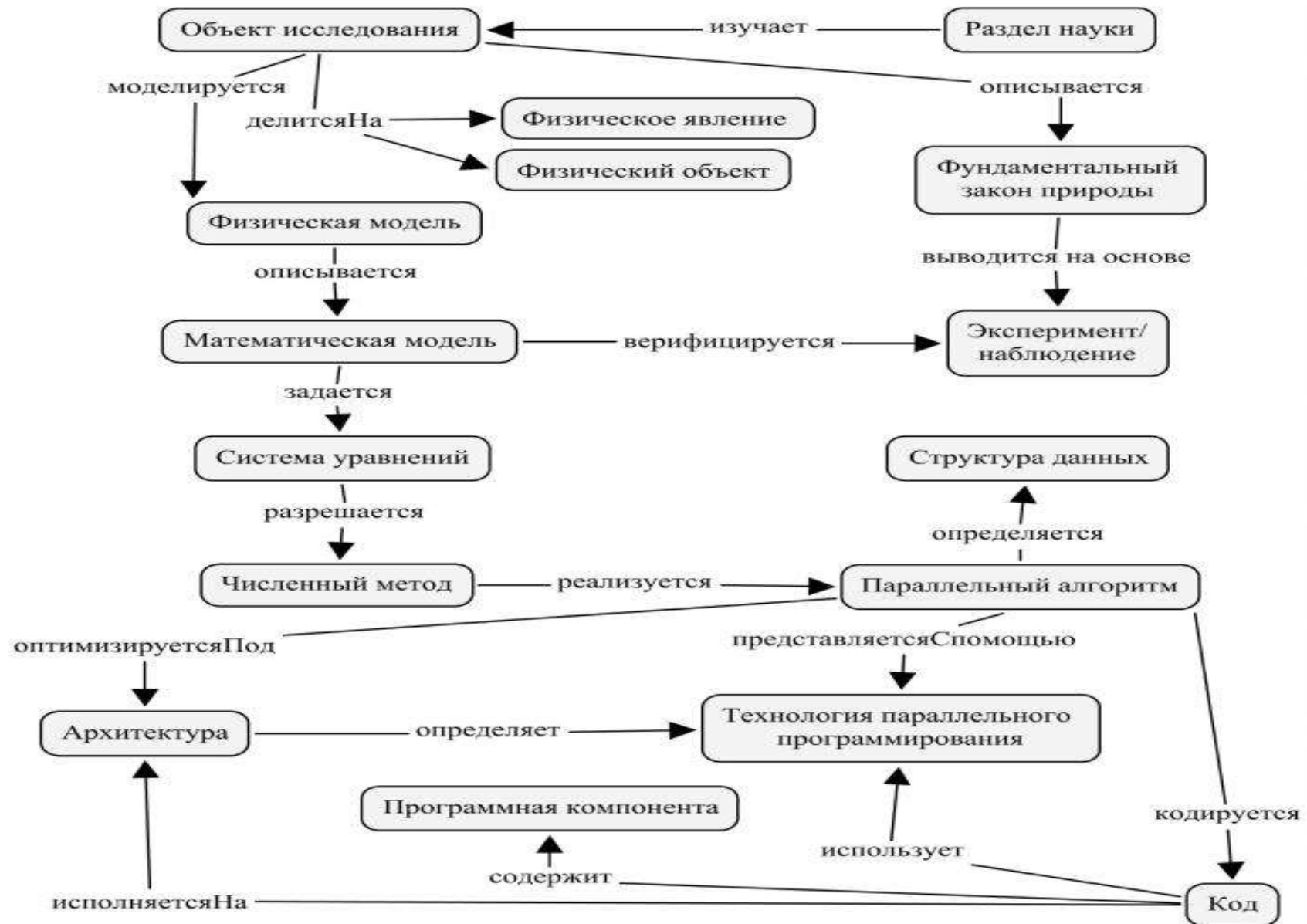
КОНЦЕПТУАЛЬНАЯ СХЕМА ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ



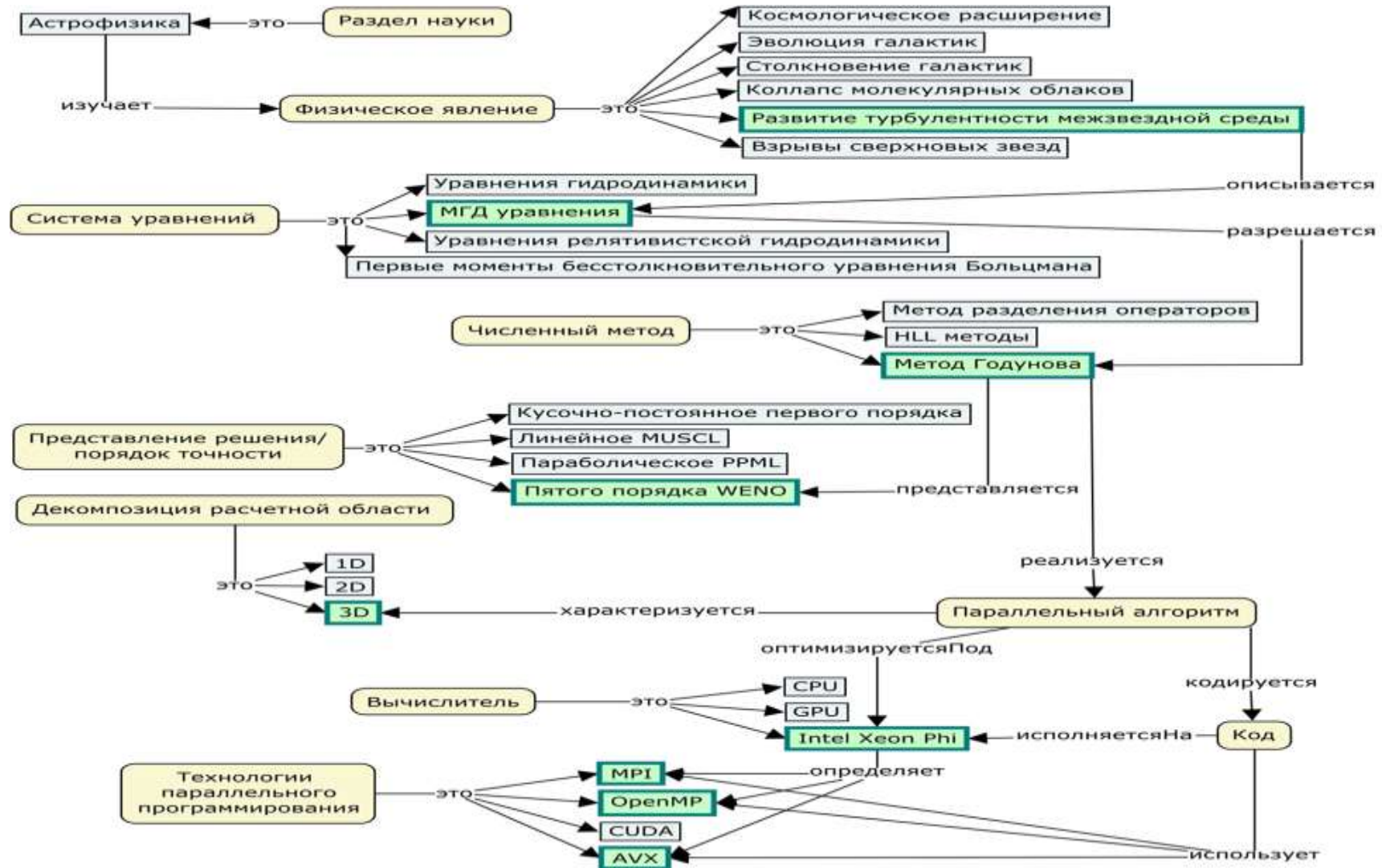
КОМПОНЕНТЫ СИСТЕМЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ

- Интеллектуальная поддержка обеспечивается следующими компонентами: базой знаний, экспертной системой; библиотекой программных компонент и модулем имитационного моделирования.
- Центральное звено - база знаний, включающая онтологию верхнего уровня «Решение вычислительно сложных задач математической физики», онтологию вычислительных методов и онтологию параллельных архитектур и технологий, и правила вывода, расширяющие логику этих онтологий. Используем информационно-аналитический интернет-ресурс по поддержке решения вычислительно сложных задач на суперкомпьютерах (ИАИР).
- Экспертная система предназначена для помощи пользователю в построении параллельного кода, решающего его задачу на суперкомпьютере. Содержит: машину вывода; схему решения задачи; модуль генерации кода; параллельный код. Пользователь подает на вход ЭС спецификацию задачи, а на выходе получает схему и параллельный код. Генератор кода подставляет в схему решения задачи соответствующие фрагменты кода из библиотеки программных компонент. Если нет, то их необходимо написать.
- Модуль имитационного моделирования выполняет оценку масштабируемости вычислительного алгоритма для заданной архитектуры суперкомпьютера.

ОНТОЛОГИЯ «РЕШЕНИЕ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ»



ОНТОЛОГИЯ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ АСТРОФИЗИКИ С ВЫДЕЛЕННОЙ СХЕМОЙ РЕШЕНИЯ РЕАЛЬНОЙ ЗАДАЧИ.



ПРИМЕР ПОСТРОЕНИЯ СИСТЕМЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ РЕШЕНИЯ ЗАДАЧ АСТРОФИЗИКИ

- Задача: смоделировать развитие гидродинамической турбулентности межзвездной среды на ранней стадии эволюции в максимальном разрешении для качественного определения фрагментации газа. Использовать вычислительные узлы, оснащенные многоядерными процессорами Intel Xeon Phi (например, узлы, входящие в состав Сибирского Суперкомпьютерного Центра ИВМиМГ СО РАН). Это исходная спецификация задачи.
- Выбранная архитектура Вычислителя в данном случае однозначно определяет Технологии параллельного программирования: MPI для обменов между узлами, OpenMP для задействования многоядерности и расширение AVX для векторизации кода. Таким образом, определена схема решения задачи. Модуль автоматизированной генерации кода позволит на выходе получить масштабируемый параллельный код, численно решающий поставленную задачу на выбранном типе Вычислителя с высоким порядком точности.
- Далее просчитывая задачу на небольшом количестве вычислительных узлов кластера определяем временные задержки для имитационной модели. Вводим полученные данные в систему имитационного моделирования AGNES и в модельном времени определяем оптимальное количество вычислительных ядер для окончательного расчета задачи на выбранной архитектуре суперкомпьютера.

ПОЛЬЗОВАТЕЛЬСКИЙ ИНТЕРФЕЙС ИНТЕРНЕТ РЕСУРСА

The screenshot displays a web application interface for a resource titled "Схема Годунова" (Godunov's Scheme). The browser address bar shows the URL: `localhost/tssc/index.php?r=ontology%2Fshow-individual&individual_id=http%3A%2F%2Fw...`. The page features a navigation menu with links: Главная, Онтология, and О ресурсе. A search icon and a "Вход" (Login) button are also present.

The main content area is titled "ПОДДЕРЖКА РЕШЕНИЯ ВЫЧИСЛИТЕЛЬНО СЛОЖНЫХ ЗАДАЧ НА СУПЕРКОМПЬЮТЕРАХ" (SUPPORT FOR SOLVING COMPUTATIONALLY COMPLEX TASKS ON SUPERCOMPUTERS). Below this, there are two tabs: "Табличное представление" (Table representation) and "Графовое представление" (Graph representation). The "Табличное представление" tab is active.

The interface is divided into two main sections: "Свойства объекта" (Object properties) and "Связи объекта" (Object relationships). The "Свойства объекта" section contains a table with the following data:

Название	Свойства объекта
Схема Годунова	

The "Связи объекта" section displays a list of relationships for the object "Схема Годунова". The relationships are categorized into three groups:

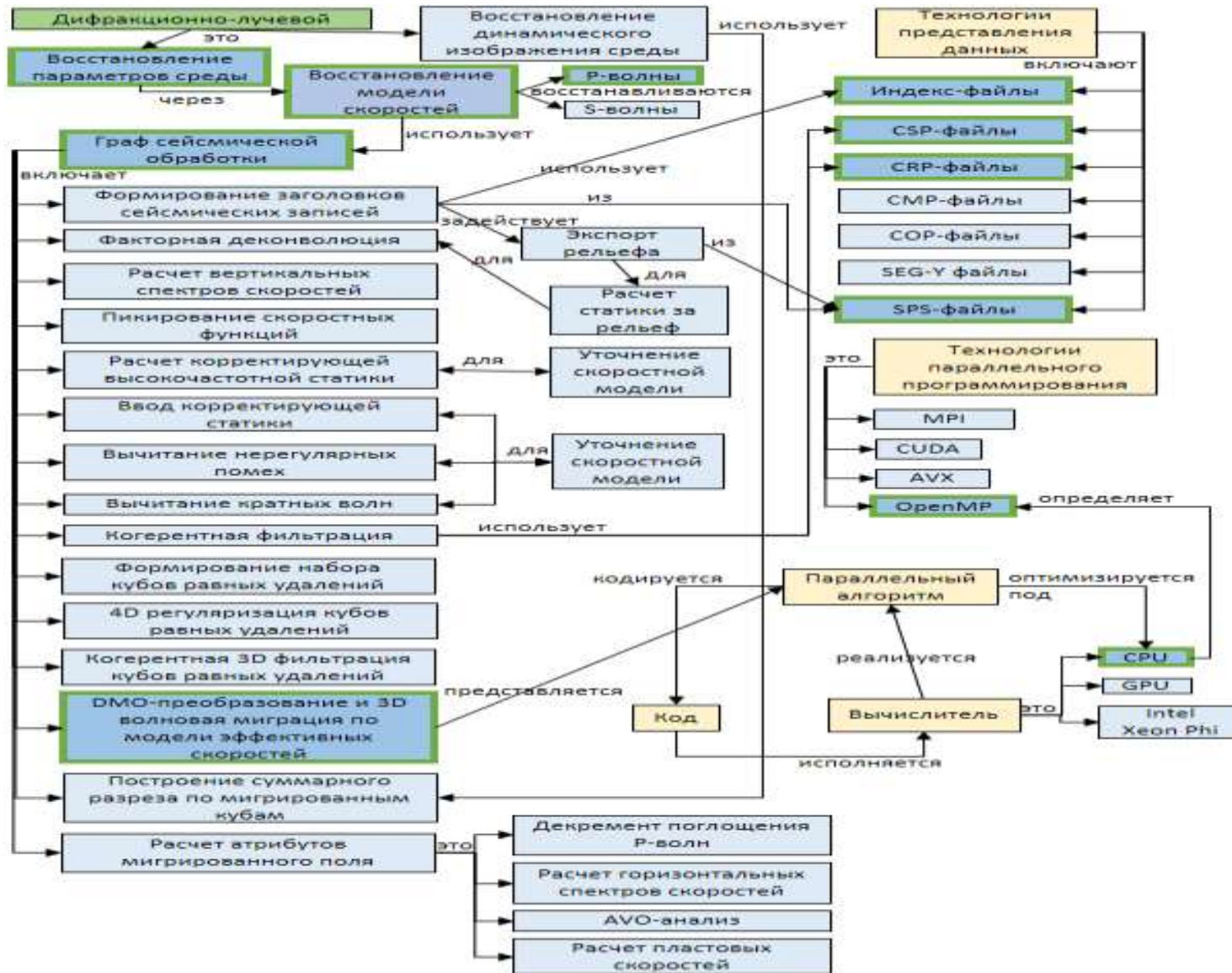
- кодируется** (is coded by):
 - Код
- оптимизируется под** (is optimized for):
 - Элемент архитектуры
- характеризуется** (is characterized by):
 - Демонстрация расчётной области

Below these, there is a section for "Обратные связи объекта" (Reverse relationships of the object), which includes:

- реализуется** (is realized by):
 - Численный метод

The footer of the page contains the copyright information: "© ИВМММГ, ИСИ 2019" and the project funding information: "Ресурс разработан при финансовой поддержке РФФИ (проект №19-07-00085)".

ОНТОЛОГИЯ ДИФРАКЦИОННО-ЛУЧЕВОГО МЕТОДА ОБРАБОТКИ СЕЙСМИЧЕСКИХ НАБЛЮДЕНИЙ



ЗАКЛЮЧЕНИЕ

- Предложена методика разработки алгоритмического и программного обеспечения для суперкомпьютеров, содержащая три связанных этапа: со-дизайн; имитационное моделирование алгоритма для исследования масштабируемости; оценка энергоэффективности алгоритма.
- Подход апробирован на трех задачах: сейсмика, астрофизики и физики плазмы.
- Исследование масштабируемости алгоритмов проводилось с применением имитационной системы AGNES, путем его отображения на гипотетическую архитектуру суперкомпьютера.
- Данный подход позволил получить для указанных задач распараллеленные коды с хорошей балансировкой и минимальным временем выполнения MPI операций.
- Показано, что алгоритмы решения данных задач для заданной архитектуры суперкомпьютера эффективно могут использовать около 1 млн вычислительных ядер.
- Разработана концепция интеллектуальной поддержки вычислительно сложных задач.
- Система интеллектуальной поддержки апробирована на задачах астрофизики и геофизики.

ПУБЛИКАЦИИ

- Glinsky BM, Marchenko MA, Rodionov AS, Karavaev DA, Podkorytov DI. Mappings of parallel algorithms on supercomputers with exaflops performance on the basis of simulation // Machine Learning and Data Analysis. Springer. 2014. V. 1, № 10. Pp. 1451 - 1465
- Glinskiy, B.M., Kulikov, I.M., Snytnikov, A.V., Romanenko, A.A., Chernykh, I.G., Vshivkov, V.A. Co-design of parallel numerical methods for plasma physics and astrophysics // Supercomputing Frontiers and Innovations Volume 1, Issue 3, 2014, Pages 88-98
- I.M. Kulikov, I.G. Chernykh, A.V. Snytnikov, B.M. Glinskiy, A.V. Tutukov. AstroPhi: A code for complex simulation of the dynamics of astrophysical objects using hybrid supercomputers // Computer Physics Communications, 2015, V. 186, P.71-80
- Glinskiy, B., Sapetina, A., Martynov, V., Weins, D., Chernykh, I. The Hybrid-Cluster Multilevel Approach to Solving the Elastic Wave Propagation Problem // Communications in Computer and Information Science, vol. 753, pp. 261-274, 2017
- G. Zagorulko, Y. Zagorulko, B. Glinskiy, A. Sapetina. (2019) Ontological Approach to Providing Intelligent Support for Solving Compute-Intensive Problems on Supercomputers. In: Kuznetsov S., Panov A. (eds) Artificial Intelligence. RCAI 2019. Communications in Computer and Information Science, vol 1093, pp 363-375. DOI: 10.1007/978-3-030-30763-9_30
- D. Vins, B. Glinskiy, I. Chernykh. Analysis of Means of Simulation Modeling of Parallel Algorithms // RuSCDays 2018, CCIS 965, pp. 1–11, 2019
- B. Glinskiy, Y. Zagorulko, G. Zagorulko, I. Kulikov, A. Sapetina. The Creation of Intelligent Support Methods for Solving Mathematical Physics Problems on Supercomputers. Russian Supercomputing Days 2019, Springer International Publishing 2019, 427-438, DOI 10.1007/978-3-030-36592-9_35

**Спасибо
за внимание!**
gbm&sscc.ru
<http://www.sscc.icmmg.nsc.ru/>