

Концепция интегрированного вычислительного
окружения для решения междисциплинарных
прямых и обратных задач математического
моделирования

В. П. Ильин

Институт вычислительной математики
и математической геофизики
Новосибирский государственный университет

Международная конференция
“Суперкомпьютерные дни в России”

Москва
24–25 сентября 2018 г.

Содержание

- Миссия математического моделирования
- Проблемы и архитектурные тенденции прикладного программного обеспечения (ППО)
- Функциональное наполнение методологических стадий и структура базовой системы математического моделирования
- Технические требования к созданию ППО
- Проблемы и подходы распараллеливания алгоритмов
- Компонентные принципы и технологии программирования
- Научно-организационные и инфраструктурные вопросы

Миссия математического моделирования

- вызовы суперкомпьютерной эволюции

-драматический рост мощностей суперкомпьютеров (закон Мура); - $10^8 - 10^{10}$ устройств арифметики и памяти - переход количества в качество (отказ от программных ограничений на число степеней свободы и количество вычислительных устройств, Intensive Data Computing);
- гетерогенные кластеры (GPU, PHI) и многоядерные процессоры с распределенной и иерархической общей памятью, минимизация энергозатратных коммуникаций (100мВт->20 мВт) ;

- суперзадачи пост-индустриальной эры

- междисциплинарные прямые и обратные задачи с реальными многомасштабными геометрическими структурами, контрастными материальными свойствами, нелинейной реологией и недоопределенными данными; - требования предсказательного моделирования с высоким разрешением расчетных данных, подтверждаемых натурными измерениями и технологиями валидации численных результатов;

- **новая вычислительная математика и нео-информатика**

- активное проникновение новых теоретических открытий в вычислительную математику (внешние дифференциальные и дискретные формы, симплектическая геометрия гамильтоновых систем, методы теории групп и графов, вычислительная топология и т.д.);

- высокопроизводительные вычислительно-информационные технологии масштабируемого параллелизма и гибридного программирования;

автоматическое построение математических моделей и алгоритмов;

- фабрика языков и компиляторов;

- искусственный интеллект - от палео-информатики к нео-информатике;

- **новый путь познания, наряду с теоретическими и экспериментальными изысканиями**

от фундаментальных исследований до оптимизации производственных процессов;

- **глобализация математического моделирования**

смена парадигмы ученых-теоретиков и экспериментаторов, появление новых массовых профессий—конечных пользователей технологий моделирования (моделлеров) в широких производственных, экономических и социальных сферах, а также в области национальной безопасности; качественный рост производительности труда и валового внутреннего продукта (ВВП).

Общесистемные проблемы и архитектурные тенденции прикладного программного обеспечения (ППО)

- классификация отраслевых задач, математических моделей и алгоритмов

Проблемы математического моделирования как общесистемной категории могут быть классифицированы по трем системам координат, практически не связанных друг с другом, в зависимости от профессиональной точки зрения заинтересованного субъекта. - *Отраслевая классификация* определяется конечной сферой приложений, и здесь могут быть названы самые разные области человеческой деятельности: энергетика, машиностроение, материаловедение, нанотехнологии, электроника, химические технологии, природопользование, медицина, социология, экономика и т.д. Приведенные наименования являются слишком общими и естественно, требуют своих внутренних спецификаций. Например, электроника бывает вакуумная или твердотельная, силовая или микро-, или СВЧ-электроника, фотоника и т.п.

Математические формулировки

- Систематизация задач моделирования по типам прикладных математических постановок в классических или обобщенных формулировках, описываемых системами дифференциальных и /или интегральных уравнений математической физики (как это традиционно называлось в XX-веке): уравнения Максвелла для электромагнитных явлений, система уравнений Навье-Стокса, описывающая гидро-газодинамические течения уравнения Ламе для упруго-пластических деформаций, уравнения Дарси для многофазных процессов в пористых средах, кинетические уравнения Больцмана для плазмы, гамильтоновы динамические системы, уравнения квантовой механики т.д. Эти же постановки могут классифицироваться по своим формальным типам в теории начально-краевых задач для дифференциальных и /или интегральных уравнений. Важно отметить, что фактически для каждой естественно-научной проблемы может формулироваться иерархическая система математических моделей, в которых по мере их усложнения учитываются все более тонкие физические или другие эффекты. Естественно, что этапы численного решения практических задач должны предваряться исследованиями теоретических свойств искомого решения (существование и единственность, корректность и устойчивость, и т.п.).

● **ВЫЧИСЛИТЕЛЬНЫЕ МОДЕЛИ, МЕТОДЫ И ТЕХНОЛОГИИ.**

Богатое разнообразие математических задач порождает еще большее многообразие методов их решения. Дискретизация непрерывных постановок осуществляется на последовательностях адаптивных сеток с локальными сгущениями в окрестностях сингулярности решений. Современные аппроксимационные подходы (методы конечных объемов и элементов, спектральные алгоритмы, численные интеграторы и др.) базируются на компактных вычислительных схемах высокого порядка точности, удовлетворяющих основным законам сохранения. Ресурсоемкие методы вычислительной алгебры реализуются с помощью высокопроизводительных параллельных технологий средствами гибридного программирования на гетерогенных вычислительных системах с распределенной и иерархической общей памятью. Решение обратных задач осуществляется алгоритмами условной минимизации целевых функционалов при заданных линейных и/или нелинейных ограничениях, на основе многократного решения параметризованных прямых постановок. Сочетание универсальности и эффективности подходов обеспечивается избыточностью вариантов алгоритмов с возможностью выбора наиболее экономичной версии, а их практическая робастность проверяется систематическим тестированием на представительных наборах методических примеров.

Обзор развития традиционных программных приложений

Прикладное программное обеспечение (ППО) с организационной точки зрения делится на общедоступное, или открытое (Open Source), и на закрытое, или коммерческое. Однако нас больше интересует его классификация по функционально-методологическим признакам, которая определяет три основные типа ППО.

- *Пакеты прикладных программ (ППП)* - это проблемно-ориентированное обеспечение конкретного класса (или классов) практических задач, включающего комплексную поддержку всех основных стадий моделирования и относительно слабо зависящее от внешних программных продуктов. Примеры высоко развитых и широко распространенных ППП-NASTRAN, ANSYS, FreeFEM, NGSolve, FlowVision.

- *Библиотеки алгоритмов и программ* представляют собой методо-ориентированное обеспечение, содержащее достаточно представительный набор методов для решения определенного математического класса задач. Примеры: "всеобъемлющая" библиотека NETLIB, библиотеки алгебраических методов MKL INTEL, EIGEN, SPARSE KIT.
- *Технологические инструментарии* по структуре могут также быть библиотечного типа, но они больше связаны с информационным обеспечением соответствующих вычислительных процессов. Примеры: генераторы сеток NETGEN, GMESH, графические пакеты PARAVIEW, VISUAL STUDIO и т.д.

В классификацию ППО мы не включаем системное программное обеспечение (СПО), а также такие самодостаточные продукты, как САПРовские разработки (CAD, CAM, CAE, PLM), имеющие свой традиционный рынок, но проявляющие тенденции интеграции с проблемами моделирования.

Концепция интегрированных вычислительных окружений (ИВО)

Идеология ИВО заключается в создании открытой инструментальной среды, ориентированной на автоматизированное построение вычислительных моделей, алгоритмов и технологий, классифицированных по всем основным стадиям моделирования широкого круга процессов и явлений.

Примерами таких проектов являются Open FOAM, DUNE (Distributed Unified Numerical Environment), MATLAB (коммерческий продукт университетской направленности, имеющий свою открытую версию OCTAVE), INMOST (разработка ИВМ РАН), а также разрабатываемая в ИВМиМГ СО РАН базовая система моделирования БСМ, рассчитанная на поддержку деятельности как разработчиков ППО, так и широкого круга конечных пользователей с разнообразной профессиональной подготовкой.

Функциональное наполнение и структура базовой системы моделирования (БСМ)

Ядро БСМ состоит из автономно разрабатываемых программных блоков, поддерживающих соответствующие стадии моделирования и взаимодействующих между собой и с внешним миром посредством согласованных структур данных. *Геометрическое и функциональное моделирование* обеспечивается модулем VORONOI, реализуемым на базе отечественной САПР-овской разработки ГЕРБАРИЙ. Содержание данной стадии заключается в описании заданной математической модели поставленной задачи для последующего автоматизированного формирования вычислительно-информационной модели, которая в итоге представляется исходным или исполняемым программным кодом. Результатом выполнения данного этапа является формирование геометрической и функциональной структур данных (ГСД и ФСД), обеспечивающих информационное согласование с типовыми САД-овскими продуктами

Генерация сеток поддерживается библиотекой DELAUNAY, которая включает ряд алгоритмов для построения специальных дискретизаций в характерных подобластях, а в целом предназначена для формирования квазиструктурированных (гибридных) сеток с возможностями использования разных типов элементов, или объемов, в различных сеточных подобластях, при активном использовании таких известных разработок, как NETGEN или GMESH. Результатом работы данного этапа является сеточная структура данных (ССД), которая в совокупности с ГСД и ФСД полностью отображает решаемую задачу на дискретном уровне. В целях обеспечения масштабируемого распараллеливания формирование сеточных подобластей должно осуществляться распределенным образом.

Аппроксимация уравнений реализуется в БСМ в рамках библиотеки CHEBYSHEV, осуществляющей поэлементное вычисление локальных матриц и сборку глобальной матрицы (при необходимости - в распределенном по MPI-процессам варианте) на разных типах сеточных ячеек для решаемых с помощью МКО или МКЭ различных порядков точности краевых задач. Исходной информацией здесь служит совокупность ГСД, ФСД и ССД, а итогом выполнения данного этапа является алгебраическая структура данных (АСД), представляющая сгенерированную систему линейных алгебраических уравнений (СЛАУ) в разреженных сжатых форматах (например, в CSR—Compressed Sparse Row).

Решение алгебраических задач в рамках БСМ поддерживается комплексом параллельных алгоритмов из библиотеки KRYLOV, представляющей собой интегрированную вычислительную среду для реализации широкого класса СЛАУ на базе современных многосеточных подходов, а также итерационных алгебро-геометрических методов декомпозиции в подпространствах Крылова, ускоряемых с помощью различных вариантов малоранговых аппроксимаций исходной матрицы. Масштабируемое распараллеливание на гетерогенных кластерах достигается средствами гибридного программирования с помощью систем MPI, OpenMP, CUDA и векторизации арифметических операций. Эффективное функционирование модулей основывается на активном использовании экономичных функций библиотеки MKL INTEL, а также допускает использование других внешних алгебраических решателей или вспомогательных вычислительных инструментов.

Технические требования к созданию ИВО

- Гибкая расширяемость состава моделей решаемых междисциплинарных проблем, а также методов решения прямых и обратных задач, без программных ограничений на число степеней свободы и на количества используемых вычислительных процессоров и/или ядер.
- Адаптируемость к эволюции компьютерных архитектур. Компонентные технологии обеспечения согласованности внутренних и внешних межмодульных интерфейсов.
- Универсальные и конвертируемые структуры данных, согласованные с имеющимися распространенными форматами и поддерживающие возможности переиспользования внешних программных продуктов

- Многоязыковость и кросс-платформенность программного функционального наполнения, открытость к согласованному участию в проекте различных групп разработчиков.
- Высокая производительность программного кода с масштабируемым параллелизмом на основе средств гибридного программирования на гетерогенных МВС с распределенной и иерархической общей памятью.
- Наличие разнообразных пользовательских интерфейсов с ориентацией на широкое применение в различных производственных сферах.

Сформулированные требования призваны обеспечить длительный жизненный цикл проекта, а также такие противоречивые качества, как эффективность, универсальность, высокую производительность и широкую востребованность.

Проблемы и алгоритмы масштабируемого распараллеливания вычислений

- Возможности грубых оценок производительности и эффективности распараллеливания вычислительных процессов, а также использования инструментов мониторинга (тьюнинга) загрузки устройств.
- Отсутствие математической модели параллельных вычислений на гетерогенной МВС, а также эффективных систем имитационного моделирования работы суперкомпьютера
- Наличие академических проектов при фактическом отсутствии производственных программных средств автоматического распараллеливания алгоритмов
- Методологии экспериментального исследования эффективности и производительности конкретных параллельных реализаций алгоритмов. Выработка практических рекомендаций

Компонентные принципы , инструменты и технологии программирования

- Автоматизация построения алгоритмов и их отображение на архитектуру ЭВМ
- Наукоемкие инструментари: Sparse Blas, MAPLE, METIS.
- Стили программирования:
модульное, сборочное, фрагментарное,
объектно-ориентированное.
- Компонентные технологии: COM/DCOM, CCA (Common Component Architecture) , SIDL (Scientific Interface Definition Language)

Цель - кардинальное повышение производительности
программирования.

Системное наполнение ИВО

- Средства автоматизации для верификации, вариации и тестирования различных типов алгоритмов, в том числе в разных внешних окружениях.
- Конфигурационное управление многоверсионных реализаций программных модулей
- Управление вычислительным процессом и организация расчетных сеансов, в том числе в облачных технологиях.
- Автоматизация построения и распараллеливания высокопроизводительных алгоритмов, их отображение на архитектуру ЭВМ.
- Конкретизация и поддержка стилей программирования: модульное, сборочное, фрагментарное, объектно ориентированное, функциональное
- Компонентные технологии для поддержки многоязыковости и кросс-платформенности проекта

Научно-организационные и инфраструктурные вопросы

Грядущая глобализация математического моделирования ставит ряд новых кардинальных проблем:

- Организация фундаментальных интеграционных исследований с целью создания и развития академического и университетского прототипа ИВО
- Образование инфраструктуры и поддержка взаимодействия академических и профессиональных коллективов по разработке производственного интегрированного вычислительного окружения для математического моделирования
- Создание инфраструктуры вычислительного сообщества разработчиков и пользователей ППО.
- Формирование системы образования и подготовки массовых профессий математического моделирования на базе суперкомпьютерной грамотности.

СПАСИБО ЗА ВНИМАНИЕ



THANKS FOR ATTENTION