

СИСТЕМА АНАЛИЗА И ВИЗУАЛИЗАЦИИ ДАННЫХ КОМПЛЕКСА ПАРАЛЛЕЛЬНЫХ ПРОГРАММ ДЛЯ РЕШЕНИЯ ЗАДАЧ ЭЛЕКТРОДИНАМИКИ

С.Л. Головков

Современные вычислительные возможности (распределенная вычислительная среда, суперкомпьютеры), выросшие требования к полноте и точности расчетов привели к существенному росту размерности и сложности решаемых задач математической физики. Это, в свою очередь, привело к росту объема числовых данных, которые необходимо анализировать прикладным специалистам.

В качестве примера можно привести комплекс параллельных программ для решения задач электродинамики [1]. Этот комплекс, разработанный в Институте прикладной математики им. М.В. Келдыша РАН, используется для численного решения трехмерных задач взаимодействия электромагнитного поля с веществом в полной трехмерной нестационарной постановке:

- математического моделирования электромагнитных полей, образующихся при взаимодействии ионизирующих излучений с веществом (эмиссии электронов за счет фото- или комптоновского рассеяния квантов излучения с поверхности преград в окружающую среду или внутренние полости технологических объектов; объемной ионизации газовых сред потоками гамма- или рентгеновского излучения);
- математического моделирования взаимодействия электромагнитного поля с веществом (дифракция на объектах со сложными геометрическими и электрофизическими параметрами; распространение импульсов электромагнитного излучения радиочастотного диапазона в проводящих средах).

Комплекс применяется для исследования помех в радиоэлектронном оборудовании, функционирующем в полях ионизирующего излучения; сопровождения экспериментов по облучению образцов на установках, продуцирующих ионизирующие излучения; исследования функционирования радиоэлектронного оборудования в условиях внешних электромагнитных помех; исследования изменений параметров электромагнитных импульсов при передаче на большие расстояния в реальной атмосфере и решения ряда других задач.

С каждым численным экспериментом, проведенным с помощью комплекса, связана база данных задачи, представляющая собой совокупность файлов, содержащих описание трехмерной геометрической модели расчетной области, описание сетки и вещества в ячейках сетки, описание структуры файлов данных и, наконец, сами файлы данных.

Файл данных представляет собой совокупность четырехмерных массивов, у которых одна ось – время – одинакова, а три пространственных оси могут быть разными. Элементы (точки) каждого отдельного массива могут быть по-разному привязаны к расчетной сетке. Например, могут лежать в центре ячеек, а могут - на границах ячеек.

Распараллеливание программ комплекса привело к существенному росту размерности решаемых задач и, как следствие, росту объема результирующих данных. Параметры типичных решаемых задач: сетка содержит $10^6 - 10^8$ узлов, число частиц достигает 10^{10} , контрольная точка в типичных сложных расчетах - до 50 GB, объем результатов расчета - до 40 GB. В отдельных расчетах эти объемы могут увеличиваться.

Для решения задачи эффективной поддержки анализа и представления больших объемов данных появилась потребность в разработке новых средств анализа данных, полученных в результате моделирования.

Принципы построения новой системы анализа вытекали из следующих соображений. Каждый элемент данных, полученных в процессе численного эксперимента, имеет привязку к геометрии расчетной области. Следовательно, метаданными, на основе которых строится доступ пользователя к расчетным данным, должно быть описание геометрии расчетной области. Отправной точкой анализа должна быть геометрическая модель расчетной области, на которой пользователь может указать некоторую подобласть и получить доступ к данным, связанным с выделенной подобластью.

Система анализа должна базироваться на следующих основных принципах:

- использование «визуального» представления области моделирования. Наличие такого представления обеспечивает пользователю возможность быстрого и интуитивно ясного доступа к произвольным совокупностям данных;
- интерактивный доступ к данным численного моделирования. Поскольку пользователям в процесс обработки результатов требуется возможность все более «тонкого» анализа данных, то процесс выделения подобласти расчетной области должен быть итеративным. Например, пользователь, получив сечение 3D области, должен иметь возможность выделить на этом сечении некоторое, вообще говоря, произвольное подмножество – многоугольник, отрезок, совокупность ячеек, обладающих некоторым свойством и т.п.

- выделив подобласть, пользователь должен иметь возможность использовать различные инструменты, стандартные или разработанные специально для каких-то случаев, для обработки информации, связанной с выделенными расчетными ячейками – визуализации, просмотра, редактирования, вычисления функционалов и т.п. Затем пользователь может вновь изменить (уточнить) выделенную подобласть и т.д.
- операции над выделенной совокупностью ячеек должны позволять пользователю получать данные в любом требуемом виде: числовом, графическом (1D, 2D, 3D графики и гистограммы) и т.п.

К настоящему моменту разработаны три варианта Системы анализа и визуализации. Первый из них является «псевдо-трехмерным»: он позволяет пользователю работать с любым ортогональным сечением расчетной области, но не способен предъявить пользователю трехмерное представление расчетной области. Этот вариант находится в опытной эксплуатации (назовем его «рабочим»).

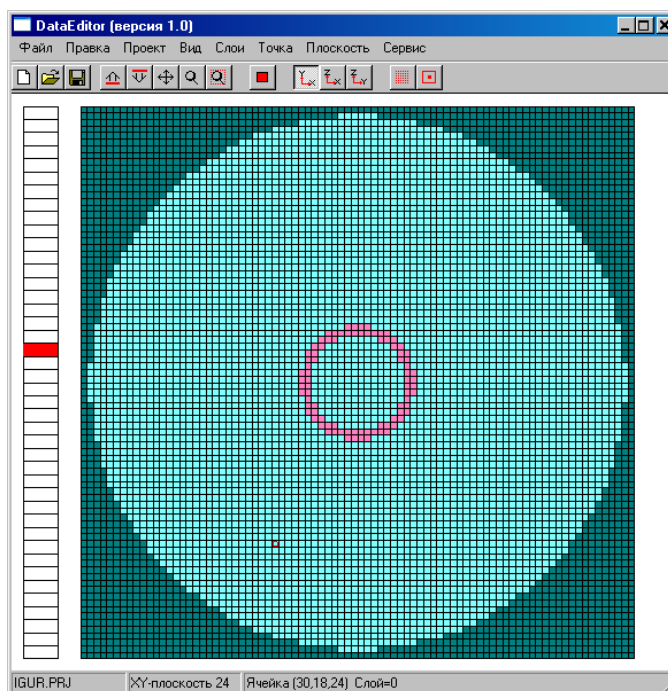
Два других варианта являются исследовательскими прототипами, на которых отрабатывается методология работы пользователя с трехмерной моделью расчетной области. Различаются они используемыми технологиями: в одном прототипе используется графическая система OpenGL, в другом - перспективная европейская интеграционная платформа SALOME (Open Source Integration Platform for Numerical Simulation) [2].

Сначала опишем общие методы использования «рабочего» варианта Системы анализа и визуализации, а затем особенности исследовательских прототипов.

СИСТЕМА АНАЛИЗА И ВИЗУАЛИЗАЦИИ («РАБОЧИЙ» ВАРИАНТ)

Система предоставляет пользователю визуальное представление расчетной области, состоящее из двух элементов:

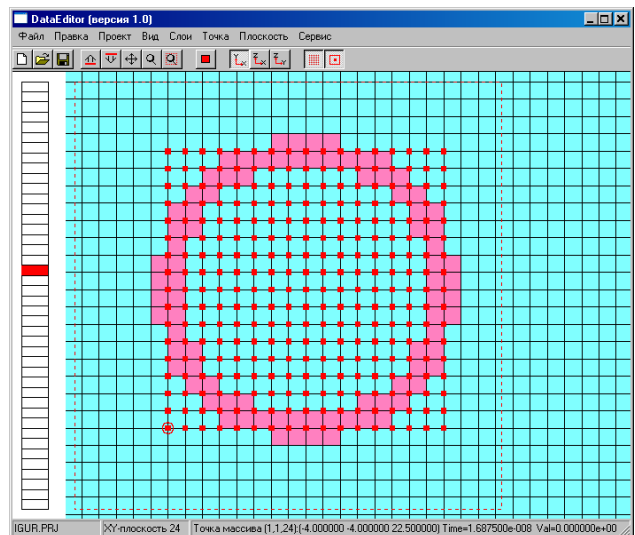
- сечение расчетной области плоскостью, перпендикулярной одной из осей координат (на скриншоте показано XY-сечение, т.е. сечение, перпендикулярное оси Z);
- «столбик» допустимых секущих плоскостей, т.е. допустимые координаты сечения (на скриншоте столбик представляет ось Z).



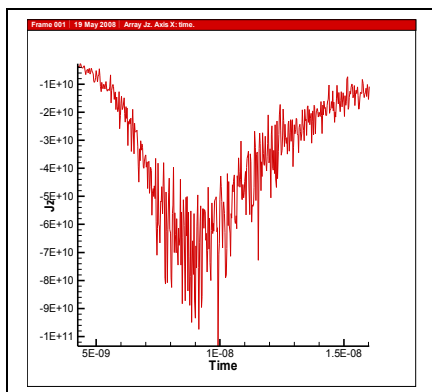
Пользователь может выбрать любую интересующую его плоскость, указав щелчком кнопки мышки тип плоскости (XY, XZ или YZ) и координату сечения на столбике плоскостей.

Пользователю предоставляются средства управления изображением (масштабирование, перемещение), средства управления цветовым представлением вещества, находящегося в ячейках, а также средства редактирования сетки (изменение содержимого ячеек сетки).

Вторым элементом управления доступом к данным является панель, позволяющая выбирать интересующие пользователя массивы и точки (или интервалы) времени. Допустим, пользователь выбрал массив Jz (в списке «массивы») и точку времени (в списке «время»). Система индицирует расположение точек выбранного массива Jz на сетке. Если пользователь щелкнет мышкой по точке массива, то Система индицирует (в строке состояния) логические и физические координаты указанной точки, а также значения массива в этой точке в указанный момент времени:

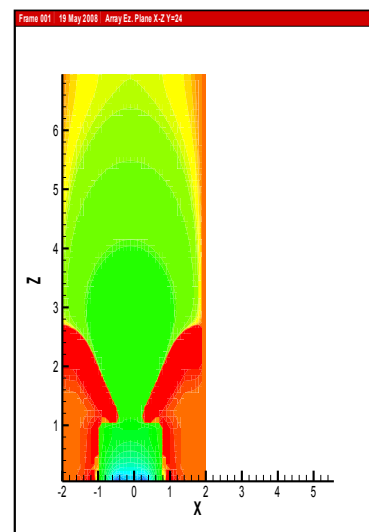
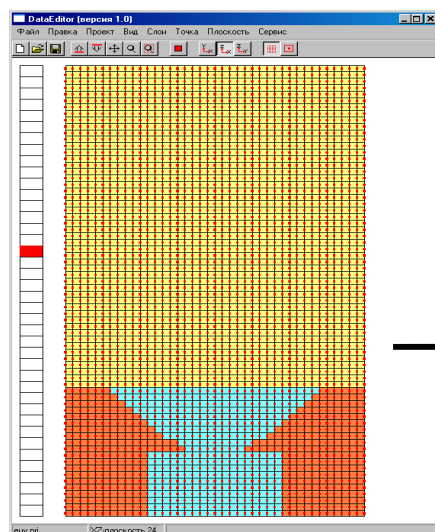


Для отдельной точки массива Система предоставляет возможность построения графика или гистограммы значений по отрезку оси времени, или выдать эти значения в числовом виде в текстовом редакторе или электронной таблице. Для этого достаточно указать мышкой точку массива, выбрать отрезок времени и выполнить соответствующую функцию меню. На экране появится окно соответствующего инструмента, содержащего искомое представление данных (например, Tecplot или WordPad):

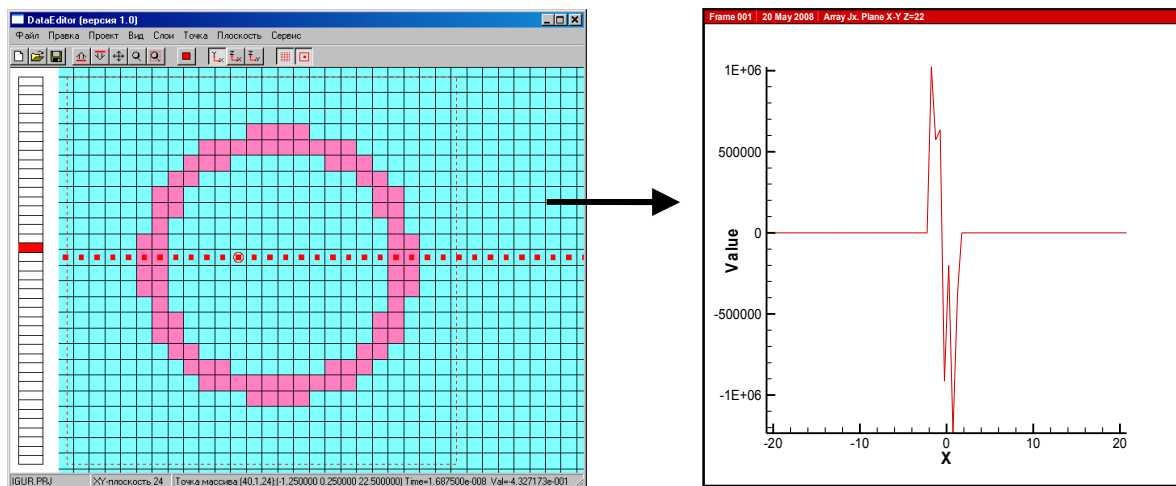


Time	Value	
1	4.250000e-009	-3.714277e+009
2	4.275000e-009	-3.527131e+009
3	4.300000e-009	-3.963983e+009
4	4.325000e-009	-4.124349e+009
5	4.350000e-009	-2.711452e+009
6	4.375000e-009	-2.797722e+009
7	4.400000e-009	-3.815496e+009
8	4.425000e-009	-4.067636e+009
9	4.450000e-009	-4.298409e+009
10	4.475000e-009	-3.973577e+009

Система анализа предоставляет возможности визуализации данных, связанных со всеми точками (или частью точек) массива, лежащими в выбранной плоскости сечения. Система определяет размерности массива на сечение и в зависимости от этого строит подходящий тип графика. Например, если сечение массива двумерное, то строится 2D (3D) график:



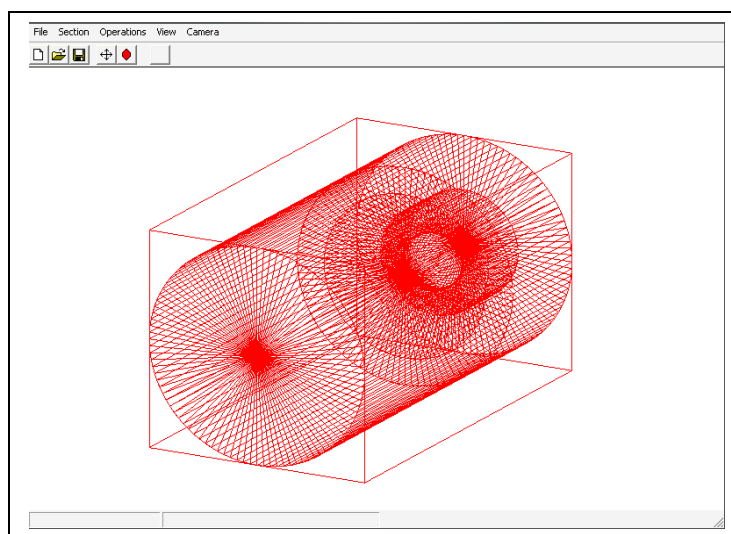
Если же сечение массива одномерное, то строится 1D график:



Система позволяет выделять часть точек массива на плоскости сечения с помощью геометрических операций (построение отрезка, прямоугольника и т.п.), строить различного типа графики, получать данные в форматах различных графических систем, объединять на одном графике значения разных совокупностей точек массива, разных точек (отрезков) времени и т.д.

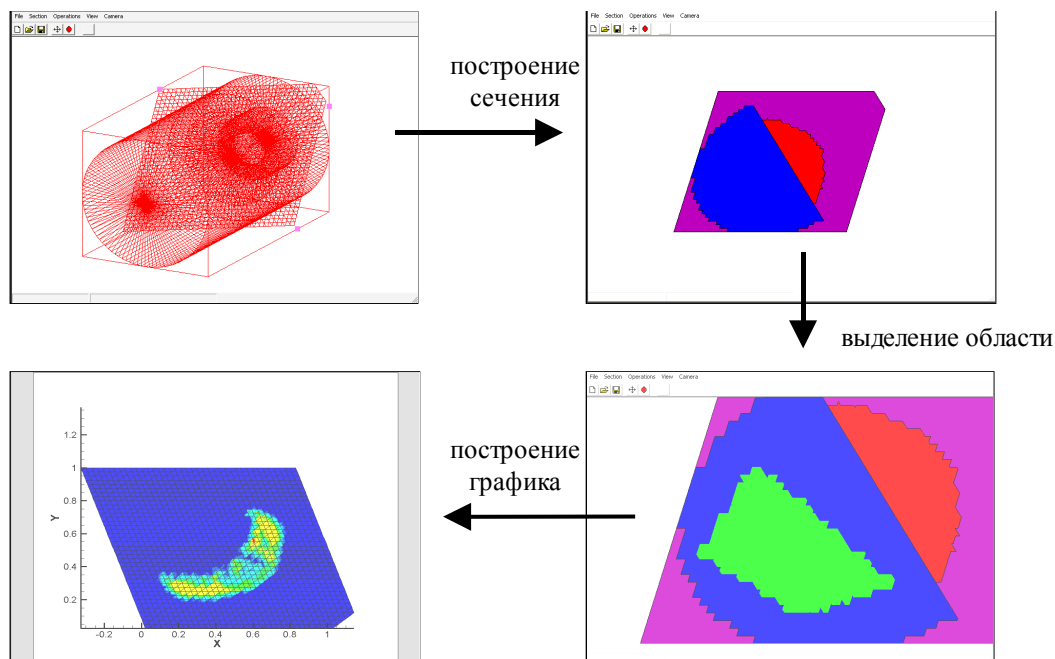
СИСТЕМА АНАЛИЗА И ВИЗУАЛИЗАЦИИ (ПРОТОТИП 1)

Рассмотрим особенности использования исследовательских прототипов. Первый прототип (разработанный на основе графической системы OpenGL) предьявляет пользователю трехмерное представление расчетной области в следующем виде:



Пользователю предоставлен обычный набор средств управления изображением: вращение, масштабирование, перемещение и т.п.

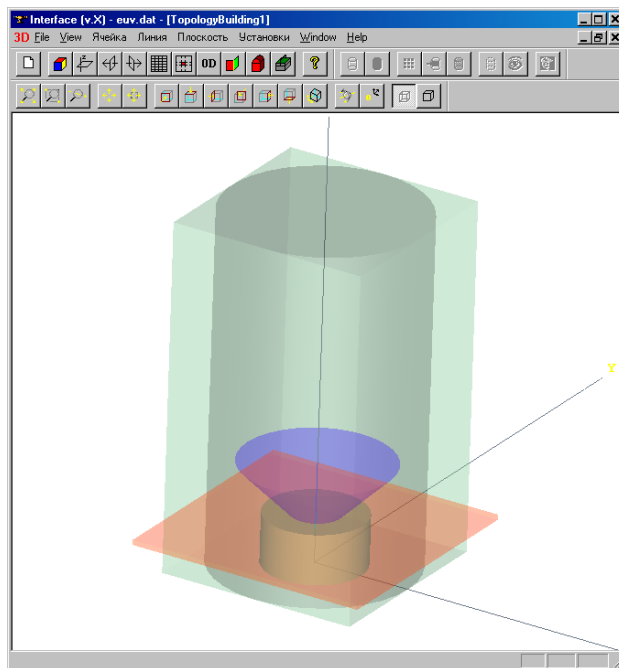
Пользователь может построить произвольное сечение области, указав на ребрах параллелепипеда точки, через которые должна проходить плоскость сечения. Затем пользователь может перейти к работе с полученным сечением: выделять на ней разнообразные области, например, выпуклые многоугольники, и получать в виде 1D, 2D или 3D графиков данные, связанные с выделенной областью:



СИСТЕМА АНАЛИЗА И ВИЗУАЛИЗАЦИИ (ПРОТОТИП 2)

В качестве технологической основы второго исследовательского прототипа была использована европейская интеграционная платформа SALOME (Open Source Integration Platform for Numerical Simulation), а точнее, один из компонентов, лежащих в основе платформы SALOME – система OpenCascade [3]. Эта свободно распространяемая система представляет собой набор библиотек и интерфейсов, предоставляющих необходимые для разрабатываемой системы функциональные возможности: 2D и 3D визуализация, работа с геометрическими объектами, нахождение производных для кривых определенного вида, нахождение расстояния между поверхностями, базовые модули построения сеток и т.п.

Этот прототип также дает пользователю возможность работы с трехмерным представлением расчетной области:



Логика работы с этим прототипом аналогична описанному выше: пользователь может строить сечения расчетной области, затем перейти к работе с полученным сечением, выделять на сечении разнообразные области и получать в числовом виде или в виде 1D, 2D или 3D графиков данные, связанные с выделенной областью.

В дальнейшем предполагается оценить прототипы с точки зрения удобства пользователя, эффективности выполнения задач представления и анализа результатов моделирования, потребления ресурсов компьютера и т.п. Затем на основе опыта эксплуатации «рабочей» версии системы принять окончательные решения относительно технологической платформы, метафоры доступа к данным, необходимого объема функциональности, интерфейсов и т.д.

Работа выполнена при финансовой поддержке Российского Фонда Фундаментальных исследований (проекты 06-07-89085-а «Исследование и разработка архитектуры хранилищ данных для поддержки решения задач математической физики»).

ЛИТЕРАТУРА:

1. А.Н. Андрианов, А.В. Березин, А.С. Воронцов, К.Н. Ефимкин, М.Б. Марков «Моделирование электромагнитных полей радиационного происхождения на многопроцессорных вычислительных системах» // Математическое моделирование, т.20, N3, 2008, 15 с.
2. SALOME Platform. www.salome-platform.org
3. Open CASCADE Technology website. www.opencascade.org