

Проект комиссии Президента
модернизации и технологическому развитию экономики России
«Создание системы подготовки высококвалифицированных кадров
в области суперкомпьютерных технологий и
специализированного программного обеспечения»

УТВЕРЖДАЮ
Председатель экспертного совета
системы НОЦ СКТ, член-корр. РАН
В.В. Воеводин

" _____ " _____ 201__ г.

Программа дисциплины

«Визуализация научных данных на суперкомпьютерах»

**«010300 -- Фундаментальная информатика
и информационные технологии»**

Разработчик: к.ф.-м.н. Корж О.В.

Рецензент: д.ф.-м.н., чл.-корр.РАН Королев Л.Н.

Москва

Учебно-методический план курса лекций «Визуализация научных данных на суперкомпьютерах» Корж О.В.

1. Аннотация курса.

Данный курс является спецкурсом для студентов 3-5 курсов. Курс читается в осеннем семестре. В конце курса сдается экзамен. Отвечает за курс кафедра автоматизации систем вычислительных комплексов факультета вычислительной математики и кибернетики МГУ имени М.В.Ломоносова. Автор программ ассистент кафедры АСВК Корж О.В. Лекторы 2011/12 учебного года ассистент Корж О.В. и инженер Андреев Д.Ю. При чтении лекций используются компьютерные презентации.

2. Место дисциплины в учебном плане

В результате изучения дисциплины студент должен:

знать теоретические основы процесса параллельной визуализации научных данных, основные методы применения визуализации в программах высокопроизводительного моделирования,

уметь разрабатывать параллельные программы визуализации для высокопроизводительных вычислительных систем, настраивать их эффективность разрабатываемых параллельных программ с учетом особенностей целевых высокопроизводительных систем,

владеть программными средствами и технологиями для визуализации на высокопроизводительных системах.

3. Содержание дисциплины

Визуализация данных является неотъемлемой частью анализа результатов вычислительных экспериментов во многих научных областях. При реализации крупномасштабных вычислительных экспериментов на суперкомпьютерах визуализация имеет ряд особенностей. В частности, размер анализируемых данных может измеряться сотнями терабайт. В таких экспериментах для визуализации требуется использовать специализированные подходы и программные продукты для построения визуального ряда. В курсе будут рассмотрены теоретические основы визуализации научных данных, методы организации данных при визуализации крупномасштабного вычислительного эксперимента, существующие и перспективные программные средства визуализации научных данных на суперкомпьютерах. Основная задача курса – ознакомить слушателей с основами построения визуальных отображений в суперкомпьютерном моделировании, с методами и алгоритмами визуального представления данных, с программными инструментами, позволяющими быстро и высококачественно выполнять визуализацию при вычислениях на суперкомпьютерах.

4. Перечень разделов курса

Суперкомпьютерное моделирование в научных исследованиях.

Обсуждаются вопросы применения суперкомпьютеров для проведения вычислительных экспериментов в различных областях исследований. Рассматриваются примеры вычислительных экспериментов, проводимых на вычислительных комплексах МГУ имени М.В.Ломоносова. Рассматриваются различные подходы к определению

визуализации. Приводится обзор технологий визуализации, показано развитие методов и программно-аппаратных средств визуализации.

Форматы представления научных данных. Методы анализа данных больших размерностей.

Рассматриваются вопросы организации, хранения и передачи данных при крупномасштабном вычислительном эксперименте. Рассмотрены часто используемые форматы представления научных данных для различных областей исследования: netcdf, hdf5, pdb, cube и т.д. Рассматриваются основные методы сокращения размерности данных при анализе: сегментация данных, кластеризация данных, методы прореживания данных, методы оптимизации структур данных, методы сжатия данных.

Анализ и визуализация данных в суперкомпьютерном моделировании. Особенности визуализации данных больших размерностей.

Раздел посвящен обзору основных методов анализа данных при суперкомпьютерном моделировании. Рассмотрены графические абстракции, используемые при визуализации научных данных, и методы их реализации в системе визуализации. Рассматриваются подходы визуализации изолиний, изоповерхностей, срезов, моделирование с помощью глифов, визуализация потоков, визуализация векторных полей, визуализация трехмерных структур данных. Формулируются основные проблемы при визуализации данных в системах с распределенным хранением результатов вычислений.

Архитектура систем визуализации на суперкомпьютерах.

Раздел посвящен описанию подходов к проектированию архитектуры систем визуализации для массивно-параллельных вычислительных комплексов. Рассматриваются особенности систем визуализации в реальном времени, т.е. систем, где визуализация происходит одновременно с вычислительным экспериментом. Рассматривается трехуровневая иерархическая система визуализации, соответствующая архитектуре современных многопроцессорных комплексов. Рассматривается метод синхронизации через промежуточные файлы.

Методы параллельного построения изображений. Объемное построение изображений (volume rendering).

Интерактивные и неинтерактивные методы параллельного построения изображений. Методы Sort-first, Sort-last, Pixel Decompositions. Подходы к представлению сцены. Метод ray casting. Методы пересечения лучей с объектами сложной формы. Модели освещения сцены. Модели отражения и полупрозрачности.

Методы объемного построения изображений для суперкомпьютеров с гибридной архитектурой.

Рассматриваются методы параллельного построения изображений при визуализации с использованием гибридной модели написания программ: MPI+OpenMP.

Ввод-вывод при визуализации данных больших размерностей.

Рассматриваются основные проблемы при вводе-выводе больших объемов данных при визуализации. Рассматриваются подходы к оптимизации ввода-вывода, методы сжатия изображений при вводе-выводе.

Библиотека визуализации VTK.

Рассматриваются конвейерная визуализация, классы представления данных, цветовые схемы, методы построения изображений, методы интерактивной визуализации, методы представления данных, интерфейс. Рассматривается графическая модель построения изображений и модель объемного представления данных для построения изображений. Рассмотрены примеры использования библиотеки VTK в разных прикладных областях. Рассматриваются примеры визуализации сцен из графических примитивов, методы визуализации функций, методы визуализации параметрических кривых, методы визуализации графов, методы визуализации данных, заданных на трехмерных сетках.

Библиотека визуализации VTK: практическое занятие по использованию библиотеки VTK на суперкомпьютере.

Практическое задание предполагает выполнение трех заданий. Задание №1 предполагает запуск готовой программы для освоения основных механизмов работы библиотеки VTK. Задание № 2 предполагает самостоятельное написание программы, визуализирующей заданную сцену из графических примитивов. Задание № 3 предполагает выполнение визуализации для файла с реальными данными некоторого вычислительного эксперимента и предполагает использованием метода объемного построения изображений. Используются данные, сохраненные в одном из стандартных форматов представления научных данных.

Система визуализации ParaView

Рассматриваются конвейерная визуализация, классы представления данных, цветовые схемы, методы построения изображений, методы интерактивной визуализации, методы представления данных, интерфейс. Особое внимание уделяется распараллеливанию построения изображений на многопроцессорной системе. Рассматриваются алгоритмы параллельного построения изображений, использованные в системе ParaView. Рассматриваются примеры использования системы ParaView в разных прикладных областях.

Система визуализации ParaView: практическое занятие по использованию библиотеки ParaView на суперкомпьютере.

Практическое задание предполагает выполнение двух заданий. Задание №1 предполагает выполнение шагов по инструкции для получения визуального отображения заданных данных. Задание № 2 предполагает выполнение самостоятельной визуализации данных, заданных в некотором стандартном представлении, с использованием системы ParaView.

Система визуализации VisIt

Рассматриваются методы представления данных, цветовые схемы, методы построения изображений, методы интерактивной визуализации, методы представления данных, интерфейс. Рассматриваются возможности создания анимации, т.е. отображения динамических данных. Рассматривается возможность визуализации данных, которые хранятся распределено. Рассматриваются примеры использования системы VisIt в разных прикладных областях.

Система визуализации VisIt: практическое занятие по использованию библиотеки VisIt на суперкомпьютере.

Практическое задание предполагает выполнение двух заданий. Задание №1 предполагает выполнение шагов по инструкции для получения визуального отображения заданных данных. Задание № 2 предполагает выполнение самостоятельной визуализации данных, заданных в некотором стандартном представлении, с использованием системы ParaView.

Обзор специализированных систем визуализации для различных предметных областей.

Приводится обзор универсальных систем визуализации: рассматриваются библиотеки OpenDX, MayaVi, SVR, pVTK. Рассматриваются специализированные системы визуализации для различных прикладных областей: молекулярной визуализации, визуализации поток, визуализации сложных структур данных и т.д.

Выполнение проекта по визуализации на суперкомпьютере.

Данное занятие предполагает самостоятельное выполнение слушателями визуализации данных для научной задачи, которой они занимаются в рамках курсовых или дипломных работ. Слушатели самостоятельно выбирают формат представления данных и систему, которую они будут использовать для визуализации. Проект выполняется индивидуально или в группе из двух человек.

5. Тематический план курса.

Номер	Наименование тем и разделов	лекции	Практические занятия	Всего часов
1	Суперкомпьютерное моделирование в научных исследованиях.	1	0	1
2	Форматы представления научных данных. Методы анализа данных больших размерностей.	1	0	1
3	Анализ и визуализация данных в суперкомпьютерном моделировании. Особенности визуализации данных больших размерностей.	2	0	2
4	Архитектура систем визуализации на суперкомпьютерах.	2	0	2
5	Методы параллельного построения изображений. Объемное построение изображений (volume rendering).	4	0	4
6	Методы объемного построения изображений для суперкомпьютеров с гибридной архитектурой.	2	0	2
7	Ввод-вывод при визуализации данных больших размерностей.	2	0	2
8	Библиотека визуализации VTK	4	0	4
9	Библиотека визуализации VTK: практическое занятие по использованию библиотеки VTK на суперкомпьютере.	0	4	4
10	Система визуализации ParaView	2	0	2
11	Система визуализации ParaView: практическое занятие по использованию библиотеки ParaView на суперкомпьютере.	0	2	2
12	Система визуализации VisIt	2	0	2
13	Система визуализации VisIt: практическое занятие по использованию библиотеки VisIt на суперкомпьютере.	0	2	2
14	Обзор специализированных систем визуализации для различных предметных областей.	2	0	2
15	Выполнение проекта по визуализации на суперкомпьютере.	0	4	4
	Итого:	24	12	36

6. Литература

1. Bruce H. McCormick, Thomas A. DeFanti and Maxine D. Brown (eds.) (1987). Visualization in Scientific Computing. ACM Press.
2. Gregory M. Nielson, Hans Hagen and Heinrich Müller (1997). Scientific Visualization: Overviews, Methodologies, and Techniques. IEEE Computer Society.
3. Will Schroeder, Ken Martin, Bill Lorensen The Visualization Toolkit An Object-Oriented Approach To 3D Graphics, 4th Edition, Kitware, Inc. publishers, 2006
4. James Ahrens, Berk Geveci and Charles Law, ParaView: An End-User Tool for Large Data Visualization. In the Visualization Handbook. Edited by C.D. Hansen and C.R. Johnson. Elsevier. 2005
5. Kitware, Inc., The Visualization Toolkit User's Guide, Kitware, Inc. publishers., 2006
6. Alexandru C. Telea, Data Visualization, A K Peters, Ltd., 2007
7. VisIt User's Manual <https://wci.llnl.gov/codes/visit/manuals.html>

Список дополнительной литературы устанавливается кафедрой.