

ГРИД ТЕХНОЛОГИИ В РЕШЕНИИ ЗАДАЧ ОБРАБОТКИ ГЕОЛОГИЧЕСКИХ И ГЕОФИЗИЧЕСКИХ ДАННЫХ

**Черемисина Евгения Наумовна¹, Сеннер Александр Евгеньевич²,
Кирпичёва Елена Юрьевна³**

¹Академик РАН, д.т.н., проф., директор Института системного анализа и управления, заведующая кафедрой САУ;

ГБОУ ВПО «Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: chere@uni-dubna.ru.

²Кандидат физико-математических наук, доцент Института системного анализа и управления;

ГБОУ ВПО «Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: senner_a_e@mail.ru.

³Кандидат технических наук, доцент Института системного анализа и управления;

ГБОУ ВПО «Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;
141980, Московская обл., г. Дубна, ул. Университетская, 19;
e-mail: kirphel@mail.ru.

В статье рассматриваются вопросы возможностей распараллеливания расчетов в области геофизических и геологических исследований. Приведен пример использования ГРИД в реальном построении 3D модели среды методом статистического зондирования. Полученный опыт продемонстрировал высокую эффективность использования среды GRID в решении рассматриваемого класса задач.

Ключевые слова: ГРИД, геология, геофизика, обработка данных, моделирование.

POSSIBILITIES OF GRID APPLICATION IN PROBLEM-SOLVING TASKS OF GEOLOGICAL AND GEOPHYSICAL DATA PROCESSING

Cheremisina Evgenia¹, Senner Alexander², Kirpicheva Elena³

¹Academician of RANS, Doctor of Science in Engineering, Professor, Director of the Institute of Systems Analysis and Management, Head of «Systems Analysis and Management» department;

Dubna International University of Nature, Society and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: chere@uni-dubna.ru.

²Candidate of Science in Physics and Mathematics, associate professor of Institute of system analysis and management;

Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: senner_a_e@mail.ru.

³Candidate of Science in Engineering, associate professor of Institute of system analysis and management;

Dubna International University of Nature, Society, and Man,
Institute of system analysis and management;
141980, Dubna, Moscow reg., Universitetskaya str., 19;
e-mail: kirphel@mail.ru.

In this article questions of opportunities of parallelization of calculations in the field of geophysical and geological researches are. The example of use of GRID in real creation of a 3D model of the environment by a method of statistical sounding is given. The got experiment showed high efficiency of use of the GRID environment in the solution of a considered class of tasks.

Keywords: GRID, geology, geophysics, data processing, modeling.

Построение моделей залегания пород в земных недрах является актуальной задачей, как с точки зрения научных исследований, так и с точки зрения практической деятельности человечества.

Научный аспект решения данной задачи состоит в стремлении понимания процессов образования и эволюции нашей планеты. Практический – в прогнозировании и поиске полезных ископаемых, необходимых для нашей жизнедеятельности. Стремительный технический прогресс, наблюдаемый в течении последних двух столетий, породил громадное количество совершенно новых устройств, технологий, материалов. В десятки и сотни раз выросли объемы промышленного производства. Это привело к потребности резкого увеличения добычи как традиционных земных ресурсов (уголь, нефть, руды и т.п.) так и ископаемых практически не востребованных ранее человечеством (уран, никель, алюминий и т.п.). По оценки экспертов в настоящее время в поверхностном слое земли (глубина не более 1 километра), доступном для добычи полезных ископаемых традиционными технологиями, разработано уже более 70 процентов полезных ископаемых.

Перечисленные выше и ряд других факторов иллюстрируют актуальность построения моделей залегания пород с целью прогнозирования, детализации и оценки объемов и перспективности залежей полезных ископаемых.

Построение моделей основано на обработке данных геофизических измерений, полученных различными методами геологической разведки. Эти данные в большинстве случаев представляют собой регулярную двумерную или трехмерную сеть. В узлах сети содержатся измеренные значения геофизических параметров (например, градиент гравитации, величина магнитного поля и т.п.).

С точки зрения обработки для построения моделей залегания пород следует выделять следующие характерные особенности этих данных:

- Большой и очень большой объем данных. Так, например, сеть размера $100 \times 100 \times 20$ содержит 2×10^5 узлов, в каждом из которых определено несколько значений физических величин. Таким образом уже такая небольшая сеть содержит порядка миллиона значений. На практике встречается необходимость обработки масштабных сетей. Объем данных при этом превышает 2 Гбайта – максимально возможный объем для хранения в оперативной памяти наиболее распространенных ЭВМ архитектуры IBM PC.
- Значительную относительную погрешность измерения данных. В подавляющем большинстве случаев измерения проводятся с помощью методов косвенного измерения (например, фиксируется отраженный сигнал).
- Возможность неоднозначной интерпретации наблюдаемой аномалии. Выявленная в результате обработки неоднородность может быть вызвана как и значительной по своим размерам аномалией, расположенной глубоко от поверхности, так и небольшой аномалией, залегающей близко к поверхности.

Перечисленные выше факторы (как и целый ряд других) обуславливают при построении моделей залегания пород использование весьма сложных алгоритмов обработки данных, многократную обработку одного и того же материала при варьировании значений параметров обработки, использование статистических и итерационных методов обработки. Время отельного расчета зачастую составляет несколько часов, а в отдельных случаях может достигать десятков часов.

Появление среды ГРИД создало предпосылки существенного сокращения процесса обработки геофизической и геологической информации для определенного класса задач, в которых возможно распараллеливание процесса обработки данных.

В данной работе обсуждается два наиболее перспективных подхода распараллеливания, не требующих значительной модификации существующего программного обеспечения и дающих значительный выигрыш по времени выполнения расчета.

Характерной особенностью трехмерных сеток в области геофизики является значительно большее количество узлов по осям X и Y (плоскость XOY располагается параллельно земной поверхности) по сравнению с количеством узлов по оси Z (Z направлена вертикально вниз). Это обусловлено быстрым нарастанием относительной погрешности измерения с глубиной и приводит к малой достоверности глубинных измерений (рис. 1):

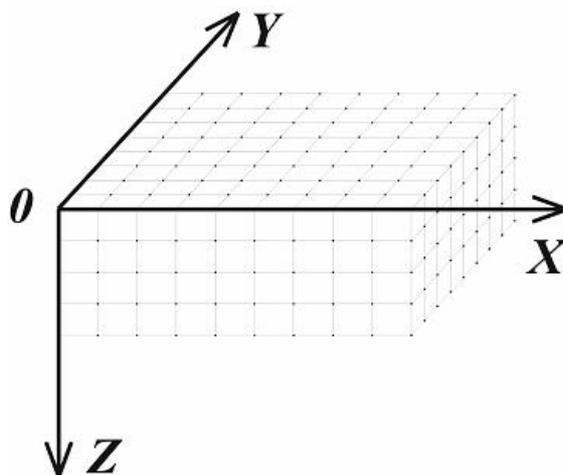


Рис. 1. Типичная сеть геофизических данных

Первый из рассматриваемых подходов состоит в следующем.

Существует обширный класс задач, алгоритмы которых основаны на вычислении в каждом узле сетки определенных величин на основании значений, хранящихся в узлах, удаленных от этого узла не далее некоторой величины R . В этом случае возможна следующая схема распараллеливания.

Исходная сетка вертикальными линиями разделяется на отдельные фрагменты параллельными плоскостями ZOX или ZOY . Алгоритм обработки, размещаемый в узлах ГРИД параметрически настраивается на обработку единственного указанного фрагмента. Обработка каждого фрагмента производится отдельным узлом ГРИД, которому передается вся сетка. Объем передач можно оптимизировать, передавая узлу ГРИД только расширенный фрагмент, состоящий непосредственно из фрагмента плюс узлов удаленных не более чем на R от граничных узлов фрагмента сетки (рис. 2).

ЭТАПЫ

Сетка исходных данных

Фрагментация сетки исходных данных

Передача фрагментов сетки в отдельные узлы ГРИД

Обработка данных в узлах ГРИД

Передача результатов расчетов узлами ГРИД

Объединение результатов

Сетка результатов расчета

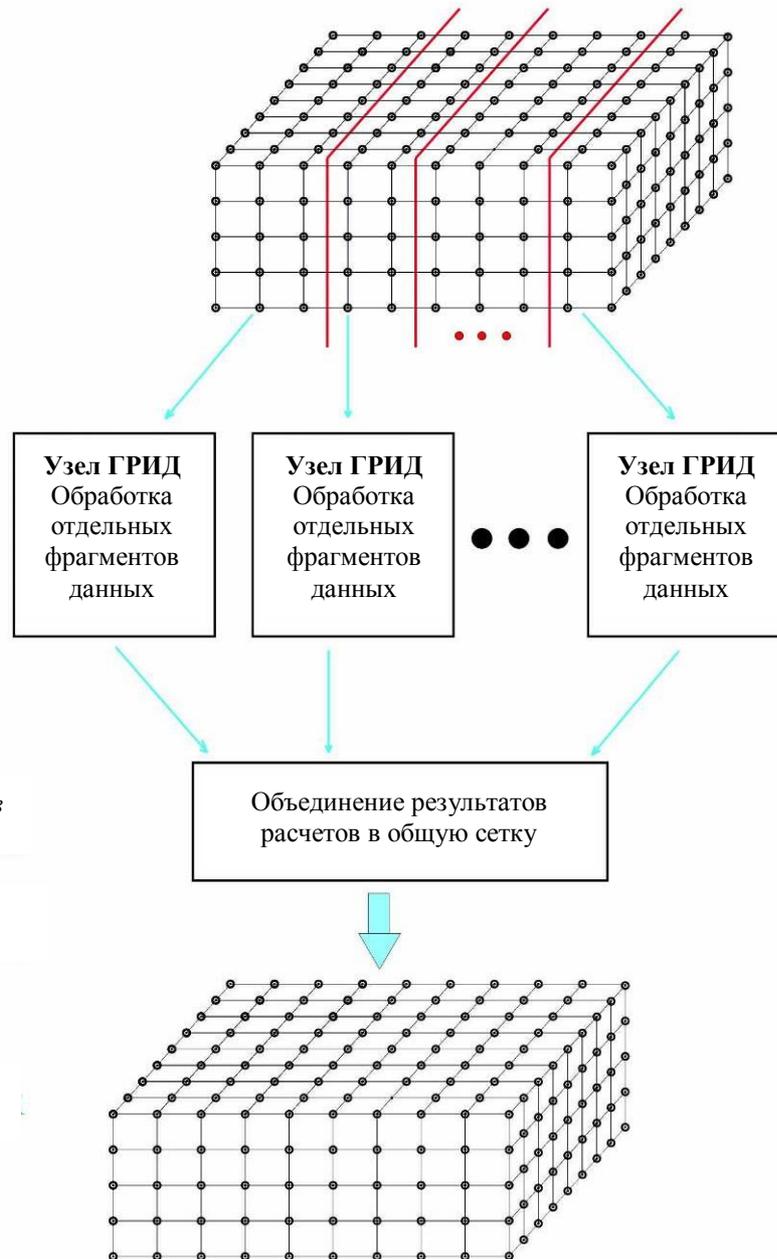


Рис. 2. Этапы обработки данных с использованием ГРИД

После завершения обработки всех фрагментов в узлах ГРИД полученные результаты интегрируются в результирующую сетку и доставляются исследователю.

Алгоритмы интеграции результатов обычно достаточно простые и не времеемкие. Поэтому оценочно можно считать, что сокращение времени расчета за счет распараллеливания пропорционально количеству фрагментов разбиения (рис. 3).

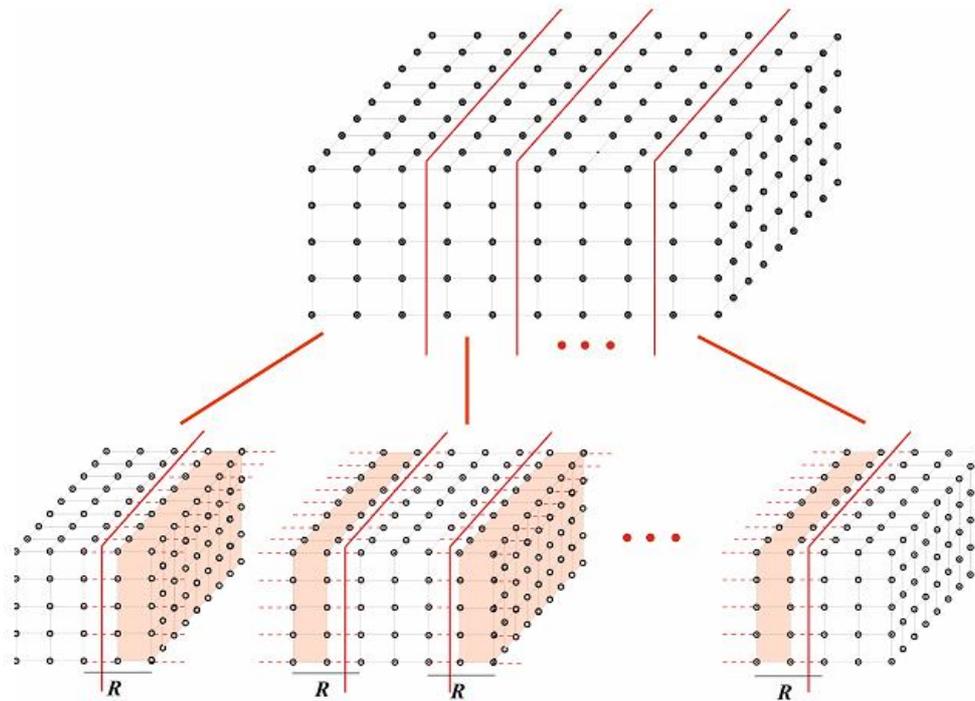


Рис. 3. Схема оптимизации передач данных по сети

Вторым подходом возможности использования распараллеливания процесса обработки является реализованный в среде GRID прикладной пакет программного обеспечения *ZondGeoStat* (статистическое зондирование) обеспечивающий на основе данных 2D сетки создание 3D модели среды.

Статистическое зондирование позволяет проследить изменение статистических характеристик, вычисленных в скользящих окнах, в зависимости от размера этих окон, что позволяет перейти к 3D объекту, несущему информацию о глубинном строении недр исследуемой территории.

Реализация процедуры расчета статистик с заданными параметрами окна, выполнена в виде независимого модуля. Это позволяет осуществить работу данного модуля на стороннем приложении. Для построения трехмерной сетки необходимо рассчитать большое количество двумерных слоев. В результате расчет требует достаточно длительного времени счета. Именно здесь актуально распараллеливание расчетов.

Методика, реализованная в пакете, основана на вычислении необходимых характеристик внутри некоторой квадратной области фиксированного размера, называемой окном. Центр каждого окна совпадает с определенным узлом сетки. Результаты расчета отдельного окна помещаются в центральный узел окна.

Отдельный проход расчета состоит в переборе всех возможных окон данного размера (рис. 4), где N_x , N_y – число узлов по оси X и Y , R – размер текущего окна.

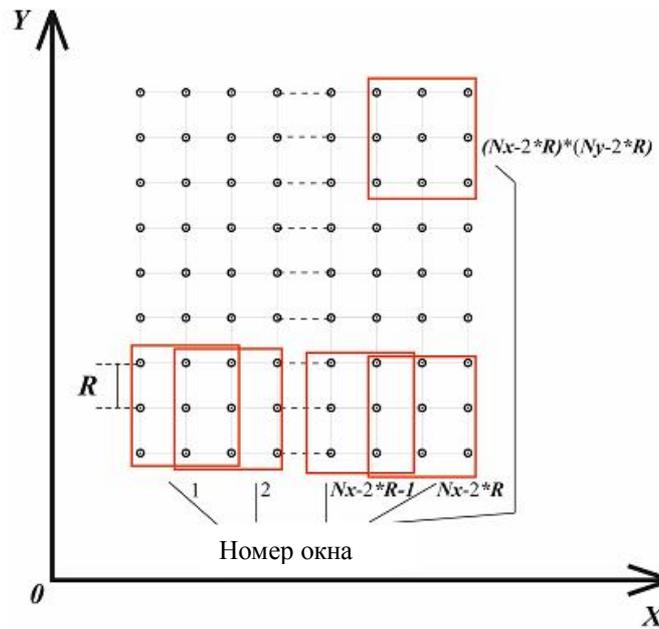


Рис. 4. Отдельный проход расчета окном размером $R = 1$

В целом расчет состоит из совокупности расчета набора окон различного размера (рис. 5).

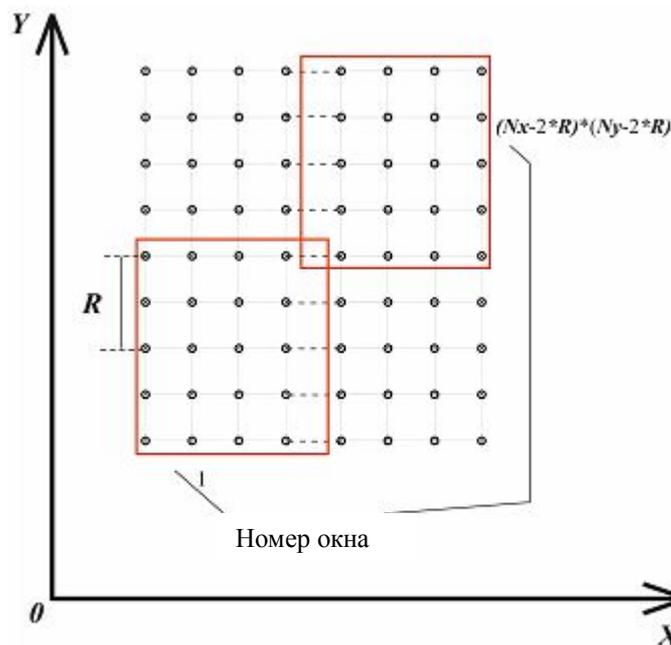


Рис. 5. Отдельный проход расчета окном размером $R = 1$

На каждом этапе расчета изменяется размер окна и процедура расчета повторяется.

Созданная в среде ГРИД реализация распараллеливает выполнение на уровне изменения окна. Исходная 2D сетка и конкретный размер окна передаются отдельному узлу ГРИД, который проводит полную обработку сетки, используя как параметр величину окна заданного размера.

По завершению вычислений узел ГРИД возвращает полученный результат. Программ интегратор объединяет результаты, полученные отдельными узлами ГРИД, и формирует суммарный результат исследования.

Типичное количество различных размеров окна при данном исследовании составляет от 10 до 30. Использование среды ГРИД, как показала практика, качественно увеличивает скорость выполнения расчета.

Подводя итог необходимо отметить, что в данной статье рассмотрены типичные, но не далеко все аспекты, возможности распараллеливания расчетов в области геофизических и геологических исследований.

Полученный опыт продемонстрировал высокую эффективность использования среды ГРИД при проведении расчетов в области геофизических и геологических исследований.

Список литературы

1. Демичев А.П., Ильин В.А., Крюков А.П. Введение в ГРИД-технологии // Препринт НИИЯФ МГУ – 2007 – 11/832. – М., 2007. URL: <http://egee.pnpi.nw.ru/doc/pp-832.pdf>.
2. Черемисина Е.Н., Никитин А.А. Геоинформационные системы и технологии. – М.: ВНИИгеосистем, 2011.