

ТЕХНОЛОГИИ В ОБРАЗОВАНИИ УНИВЕРСИТЕТ

МИКРОЭЛЕКТРОНИКА
ИННОВАЦИИ
КАТАЛИТИЧЕСКИЕ
МАТЕРИАЛЫ
ДИЗАЙН
ЛЕКАРСТВ
ТОЧКА
СБОРКИ

НАУЧНАЯ
ЛАБОРАТОРИЯ
ГЕОХИМИЯ
ИНЖИНИРИНГ
ГЕОФИЗИКА
ГИБРИДНЫЕ
МАТЕРИАЛЫ
ЭНЕРГОСБЕРЕЖЕНИЕ
НГУ

ВЫСОКИЕ
ЭНЕРГИИ
БИОТЕХНОЛОГИИ
МОДЕЛИРОВАНИЕ
НАНОТЕХНОЛОГИИ
СЕМИОТИКА

НАУКА
КОГНИТИВНЫЕ ТЕХНОЛОГИИ
МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

IT
DEEP
LEARNING
ИЗУЧЕНИЕ

МОЗГ
АРКТИКА
ЭЛЕМЕНТАРНЫЕ
ЧАСТИЦЫ
ГЕОЛОГИЯ
КВАНТОВЫЕ
ТЕХНОЛОГИИ
БИОЛОГИЯ
ТЕМНАЯ
МАТЕРИЯ
ФОТОНИКА
БИОМЕДИЦИНА
ПРИКЛАДНЫЕ
ИССЛЕДОВАНИЯ
РАЗВИТИЕ
АСТРОНОМИЯ
ГЛОБАЛЬНЫЕ ПРИОРИТЕТЫ
АСТРОФИЗИКА
БИОИНФОРМАТИКА
ЛАЗЕРНАЯ
ФИЗИКА
АРХЕОЛОГИЯ
ЭКОНОМИКА
ЗНАНИЙ
СОТРУДНИЧЕСТВО

N* Новосибирский
государственный
университет
*НАСТОЯЩАЯ НАУКА



Новосибирский государственный университет, НГУ
"09.03.01 Информатика и вычислительная техника.
Программная инженерия и компьютерные науки"

Разработка и реализация частных алгоритмов
управления распределёнными данными в
системе LuNA

Макаренко Данил Евгеньевич, студент группы 17202

Научный руководитель: Малышкин Виктор Эммануилович д.т.н.
профессор, зав. каф. ПВ ФИТ НГУ

Соруководитель: Перепёлкин Владислав Александрович ст.
преп. каф. ПВ ФИТ НГУ

Рецензент: Горкунов Евгений Владимирович, к.ф.-м.н., с.н.с. ИМ
СО РАН

23 июня 2021 г.

* Описание проблематики

Разработка эффективных параллельных программ численного моделирования на суперкомпьютерах сложна, трудоёмка и требует специальной квалификации.

Проблемы параллельного программирования:

- Декомпозиция данных и вычислений
- Распределение данных и вычислений по узлам
- Организация параллельной обработки распределенных данных
- Настройка на доступные ресурсы и обеспечение масштабируемости

* Описание проблематики

Распределение ресурсов является NP-трудной задачей, что значит, что никакой из алгоритмов распределения не дает гарантированного эффективного распределения ресурсов, подходящего для широкого класса задач, за разумное время.

* LuNA

Система фрагментированного программирования LuNA, разработанная с целью упростить написание параллельных программ, является подходящей площадкой для исследования проблемы распределения ресурсов из-за особенностей архитектуры, позволяющей сосредоточиться на разработке самих алгоритмов.

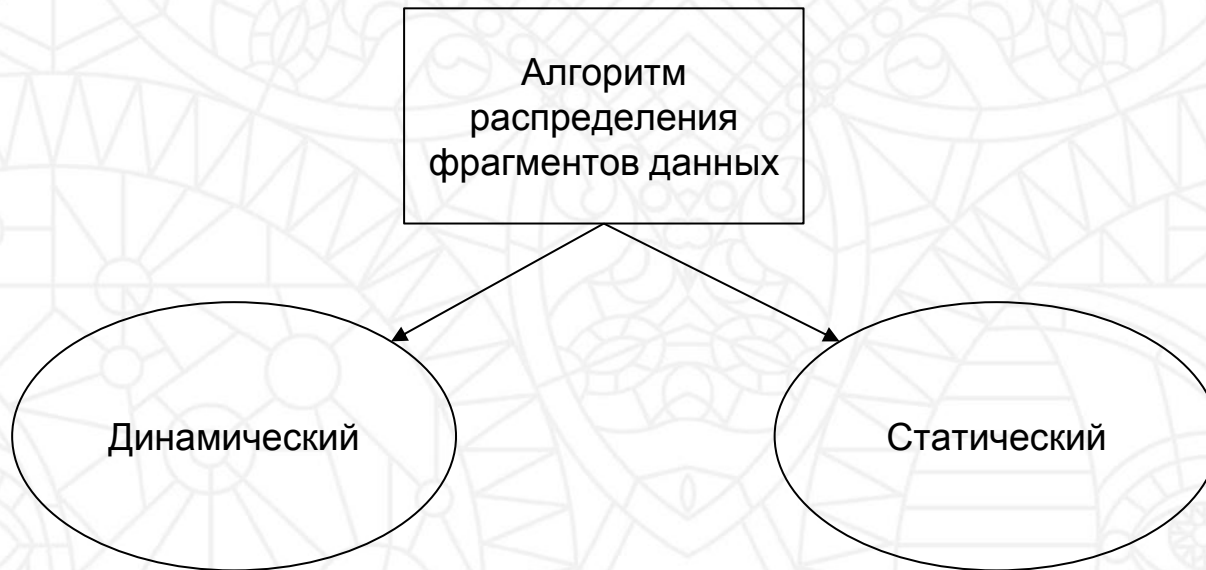
* Цель работы

Разработка алгоритма распределения фрагментов данных в системе конструирования параллельных программ LuNA.

Достижение поставленной цели требует решения следующих задач:

1. Анализ существующих алгоритмов распределения
2. Выделение основных требований для алгоритма распределения
3. Разработка и реализация алгоритма распределения фрагментов данных в системе LuNA
4. Оценка эффективности реализованного алгоритма

* Существующие решения



* Существующие решения

Алгоритмы:

1. Диффузионные алгоритмы
2. Распределение с помощью плиточных массивов
3. Алгоритм Rope-of-bids
4. Алгоритм Patch

Недостатки

1. Ограниченность кругом задач,
2. Необходимость запускать алгоритм несколько раз
3. Невозможность использования в процессе исполнения программ
4. Не масштабируемость на большое число узлов.

* Требования к алгоритму

1. Отсутствие глобальных коммуникаций
2. Обеспечение равномерного распределения
3. Уменьшения объема сообщений между вычислительными узлами
4. Децентрализованность

* Постановка задачи

Важной частью формирования программы является конструирование распределения фрагментов вычислений и данных. **Фрагмент вычислений** – независимая единица программы, которая состоит из описания входных/выходных фрагментов данных и кода (модуля, процедуры) фрагмента. Для того, чтобы сконструировать распределение нам необходимо построить отображение данных и блоков вычислений на узлы вычислительной системы. Полученное отображение должно удовлетворять следующим **критериям**:

1. Высокий уровень производительности системы при использовании полученного отображения
2. Уменьшение трудозатрат на разработку фрагментированной программы
3. Удовлетворение широкому кругу задач

Распределение можно задать статически с помощью рекомендаций. **Рекомендации** - это механизм управления системой, который при добавлении в код специальных аннотаций изменяет исполнение программы

* Предлагаемое решение



* Предлагаемое решение

Данные используются вместе, если они являются входными и выходными параметрами одного фрагмента вычислений

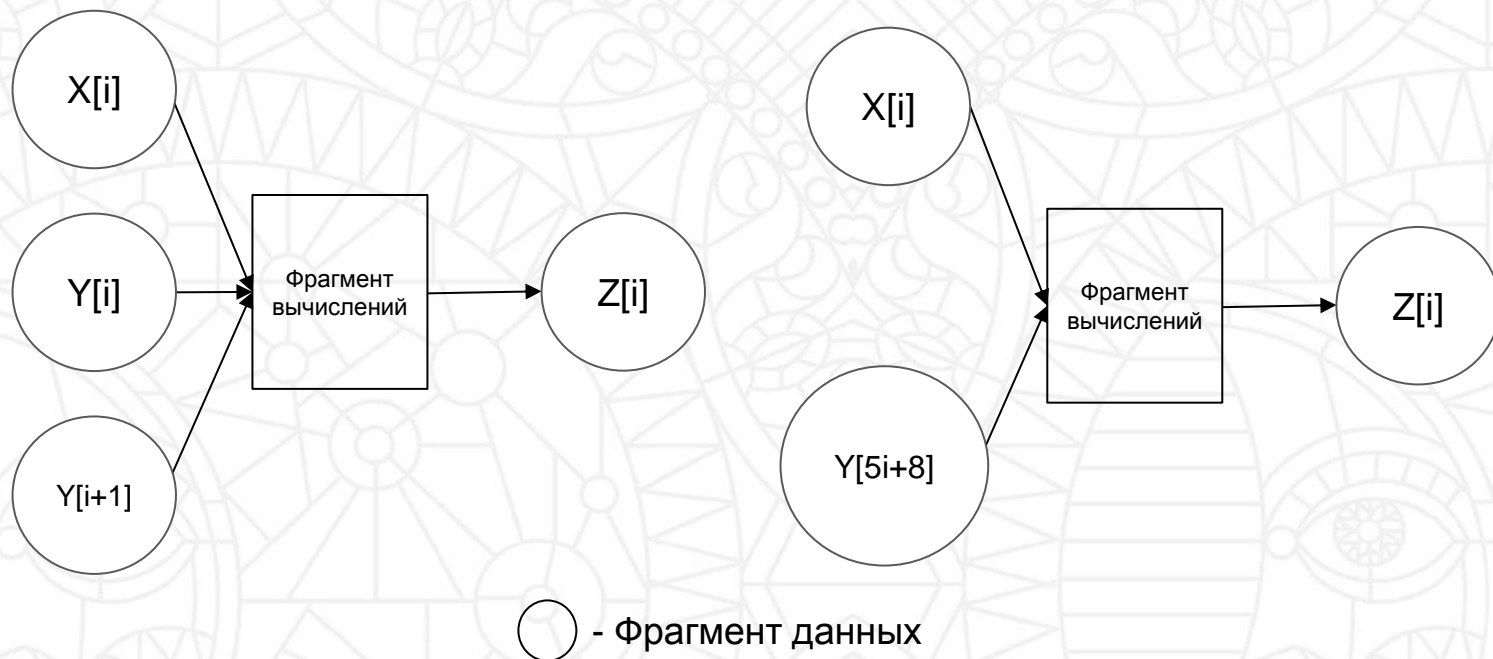


○ - Фрагмент данных

* Алгоритм

1. Выявление совместно используемых данных
обход АСД пока за обход не будет выявлено новой зависимости {
 1. Чтение аргументов функций
 2. Выявление данных имеющих статическое распределение
 3. Анализ на возможность использовать схожее распределение}
2. Применение распределения для совместно используемых данных

* Предлагаемое решение(Конфликтные ситуации)



* Результаты тестирования

Тестирование проводилось на задаче скалярного произведения двух векторов, с использованием 4 вычислительных узлов кластера. Размерность векторов равнялась 5000. Изначально, мы задали распределение для одного вектора(назовем его x)

Время работы без рекомендаций, с	Время работы с рекомендацией для $x[i],c$	Время работы с использованием алгоритма, с
30,945	27,039	23,658

* Результаты тестирования

Для проверки алгоритма на работоспособность с более сложными структурами программ была взята задача с решением уравнения Пуассона.

Время работы без алгоритма, с	Время работы с алгоритмом, с
56,396	50,248

* Заключение

На основе требований был разработан алгоритм распределения данных на основе анализа исходного кода на наличие рекомендаций для распределения данных по вычислительным узлам. Данный алгоритм является статическим и выполняется на этапе компиляции программы. Разработанный алгоритм позволяет снять с пользователя трудозатраты на расстановку рекомендаций по распределению данных, а также повышает качество конструированных программ

Защищаемые положения:

1. Разработан алгоритм распределения ресурсов по вычислительным узлам на основе статического анализа
2. реализован алгоритм распределения ресурсов в компиляторе системы LuNA
3. проведено исследование эффективности исполнения конструированных программ с разработанным алгоритмом распределения

* Публикации

Макаренко, Д. Е. Разработка алгоритма распределения фрагментов данных в системе конструирования LuNA/ Д. Е. Макаренко // Информационные технологии: Материалы 59-й Междунар. науч. студ. конф. 12–23 апреля 2021 г. / Новосиб. гос. ун-т. — Новосибирск: ИПЦ НГУ, 2021. — С. 154.

Спасибо за внимание!



РОССИЯ, 630090, НОВОСИБИРСК, УЛ. ПИРОГОВА, 2



nsuniversity.official



nsu24



@nsuniversity

WWW.NSU.RU/N/