ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ, НГУ)
Факультет информационных технологий
Кафедра параллельных вычислений
09.03.01 Информатика и вычислительная техника. Программная инженерия
и компьютерные науки

Реализация централизованного подхода к динамической балансировке вычислительной нагрузки в системе фрагментированного программирования LuNA и его сравнение с децентрализованным подходом

Выполнил: Мустафин Дамир Эркинович, студент группы 18201 ФИТ НГУ
Научный руководитель: Власенко Андрей Юрьевич, к.т.н., доц. каф. ПВ ФИТ НГУ
Рецензент: Власов Александр Александрович, к.т.н., с.н.с. ИНГГ СО РАН

Новосибирск 2022

20.06.2022

Актуальность

Эффективность исполнения программы в системе LuNA в числе прочих факторов зависит от эффективности алгоритмов балансировки нагрузки.

Существующий децентрализованный модуль динамической балансировки нагрузки:

- не полностью автоматический (необходима информация о задачах, которые следует учитывать при балансировке нагрузки);
- в некоторых случаях неэффективно использует ресурсы, следовательно программы работают дольше, чем могли бы;
- не имеет полной информации о загруженности всех узлов.

Актуально разработать модуль автоматической динамической балансировки нагрузки, лишенный указанных недостатков.

Цель

Реализация модуля централизованной динамической балансировки вычислительной нагрузки в системе LuNA и экспериментальное сравнение с модулем, использующим децентрализованный подход.

Задачи

- Обзор средств оптимизации параллельных программ.
- Формулировка требований и проектирование модуля динамической балансировки нагрузки, использующего централизованный подход.
- Реализация централизованного подхода к динамической балансировке нагрузки в системе LuNA.
- Тестирование и экспериментальное сравнение двух подходов к динамической балансировке нагрузки.

Обзор средств оптимизации параллельных программ

- MPI Tuning (подбор наилучших значений параметров).
- Intel Trace Analyzer and Collector (профилирование и анализ приложения).
- Charm++ (система параллельного программирования).
- DVM (система, создание переносимых и эффективных параллельных программ).

Некоторые из концепций, реализованных в перечисленных средствах, были адаптированы и использованы в реализации централизованного балансировщика в системе LuNA.

Подходы к балансировке нагрузки

- Статический.
- Динамический:
    - Централизованный.
    - Децентрализованный.

Система LuNA

Программа состоит из множества **фрагментов вычислений** и **фрагментов данных.**

Фрагмент вычислений (ФВ) - единица программы, которая содержит описание входных/выходных фрагментов данных и код фрагмента.

Фрагмент вычислений считается **готовым** к исполнению, если он не зависит от фрагментов данных, либо все его входные фрагменты данных уже вычислены.

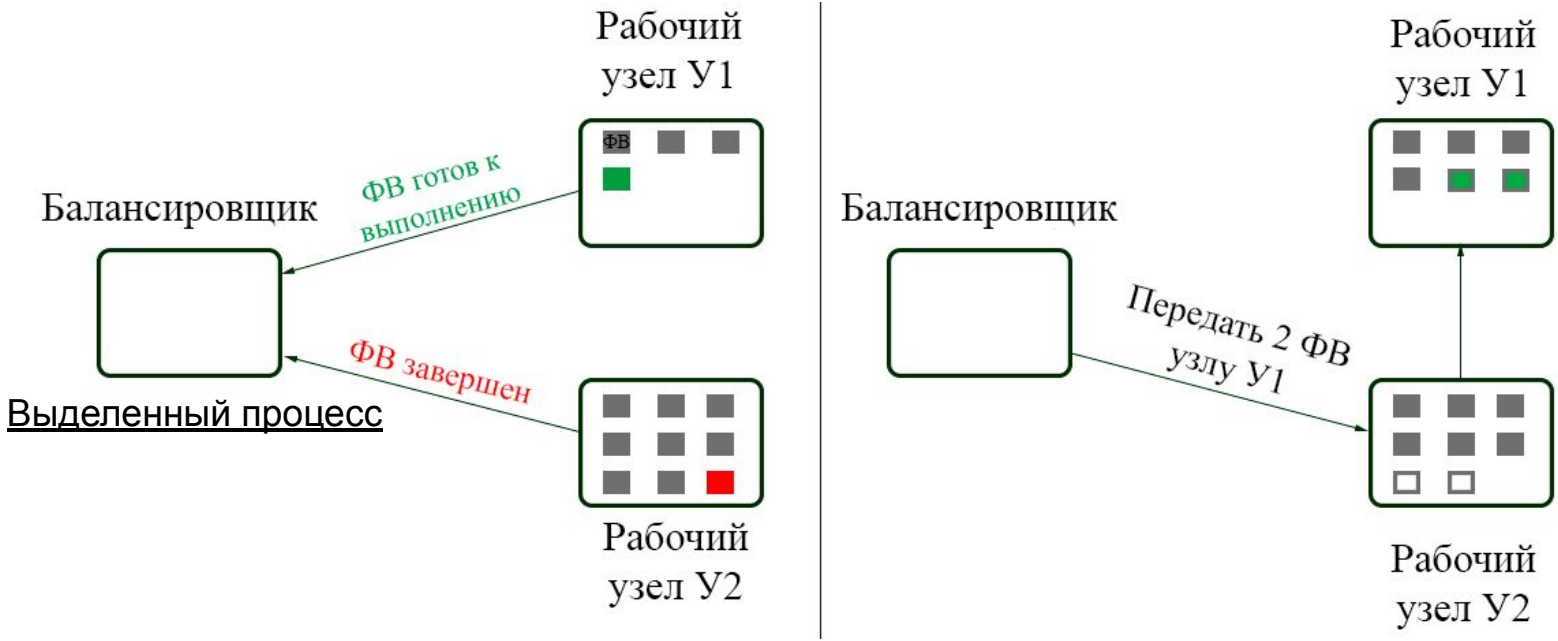
Жизненный цикл ФВ на узле:

Используемые методы и технологии

Модель взаимодействия: ведущий – ведомый.

Модуль централизованной балансировки нагрузки написан на языке C++ с использованием MPI.

Принцип работы централизованного балансировщика



Составление плана балансировки при перераспределении нагрузки

План балансировки — схема передач ФВ между узлами в процессе перераспределения нагрузки.

При составлении плана балансировки следует учитывать следующие критерии:

- **максимальной равномерности** распределения нагрузки между узлами после балансировки;
- **минимальных накладных расходов**, связанных с передачей ФВ.

фв_b^a – ФВ номер b на узле a с весом ВЕС_b^a

T_b^a – время передачи фрагментов вычислений от узла a к узлу b в процессе данной балансировки

Узел 0: $\text{фв}_0^0,\ \text{фв}_1^0, \ldots, \text{фв}_{n_0}^0$;

Узел 1: $\text{фв}_0^1,\ \text{фв}_1^1, \ldots, \text{фв}_{n_1}^1$;

. . . . .

Узел k: $\text{фв}_0^k,\ \text{фв}_1^k, \ldots, \text{фв}_{n_k}^k$.

Критерий максимальной равномерности:

$$\max_{p,q}\left(\left|\sum_{i=0}^{n_p}\text{ВЕС}_i^p - \sum_{j=0}^{n_q}\text{ВЕС}_j^q\right|\right) \to \min$$

Критерий минимальных накладных расходов, обусловленных служебными коммуникациями:

$$\min \sum_a^k \sum_b^k T_b^a \to \min$$

Определение весов ФВ на основе времени их выполнения

Группировка ФВ, вызываемых множество раз (в одну группу попадут, например, все вызовы одной и той же процедуры с разными аргументами).

Оценка весов невыполненных ФВ из одной группы средним арифметическим уже выполненных.

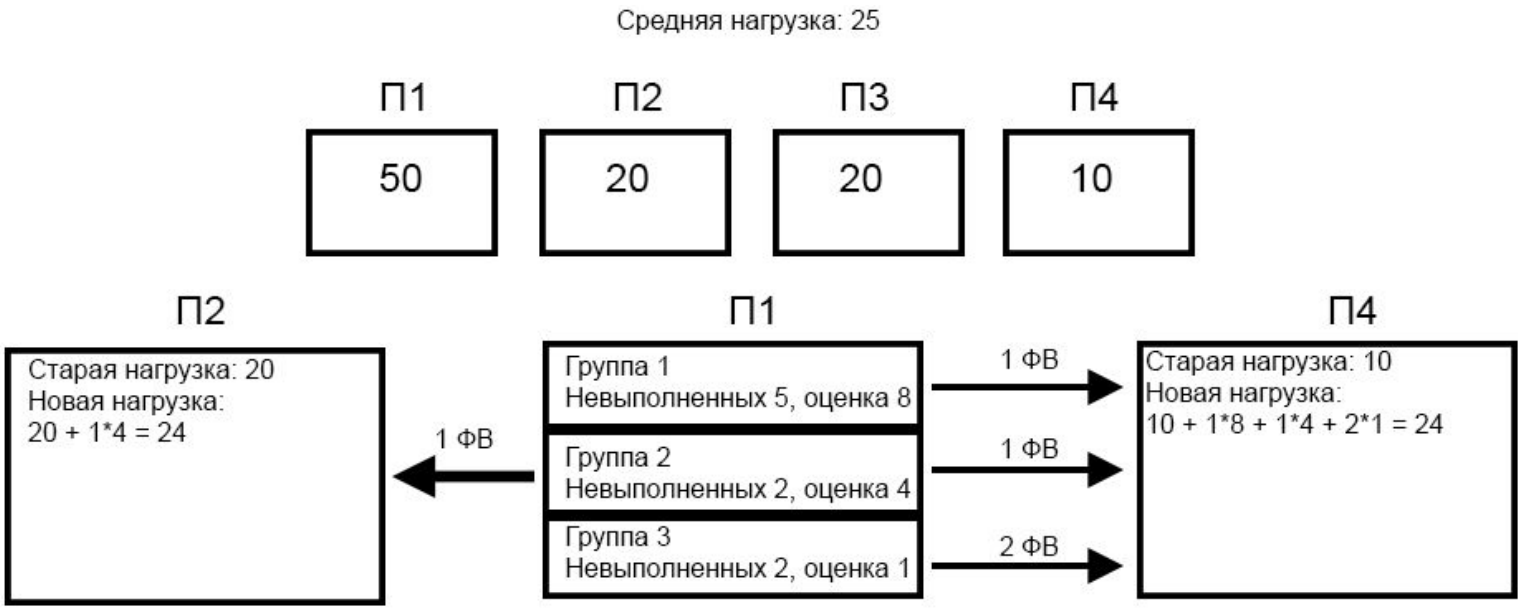
Вес ФВ - время его выполнения (в секундах).

* ФВ состоят из функций, называемых блоками. Блоки выполняются в определенном порядке.

Текущий план балансировки (итерационный)

Перемещение ФВ от самых нагруженных узлов к самым ненагруженным. В приоритетном порядке перемещаются самые тяжелые ФВ, затем менее тяжелые.

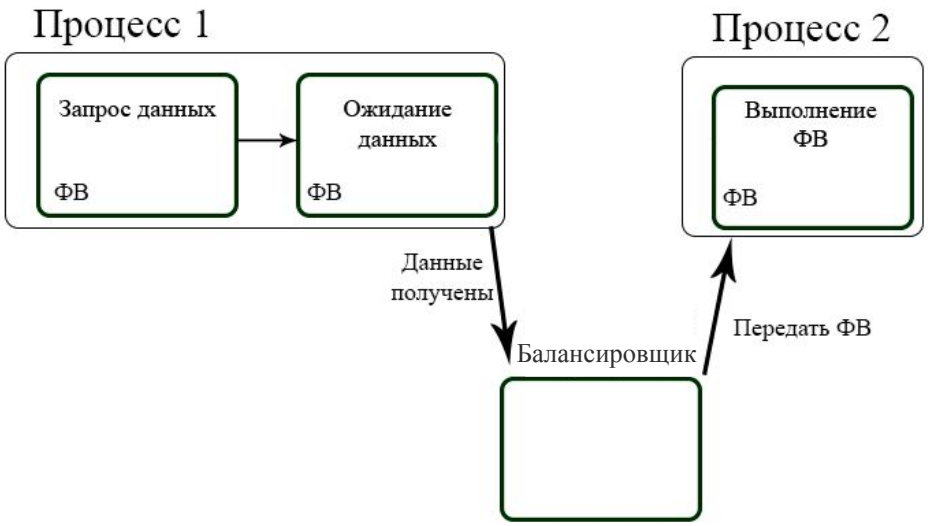
Нагрузка на узле, получающем фрагменты вычислений, не должна превышать среднюю арифметическую нагрузку по всем узлам.



Фрагменты вычислений при перераспределении нагрузки

- При перераспределении нагрузки используются только готовые к исполнению фрагменты вычислений. Они передаются вместе со своими входными данными.
- Фрагменты вычислений передаются одним сообщением.
- При передаче ФВ происходит проверка на наличие выигрыша.

Схема выполнения ФВ:



Параметры модуля балансировки нагрузки

2 параметра, влияющие на эффективность балансировки и погрешность при перераспределении нагрузки:

- JOBS_LEFT_THRESHOLD (по умолчанию 5).
  - Определяет минимальную суммарную нагрузку узлов (в секундах), при которой необходимо осуществить балансировку нагрузки.
- JOBS_DIFFERENCE_RATIO (по умолчанию 0.5).
  - Определяет минимальную разность нагрузки (в долях).
  - Если разность нагрузки между двумя узлами, равная (H1-H2)/H1, превышает данный параметр, то будет совершено перераспределение нагрузки между ними.

Тестирование ускорения и эффективности

Условия тестирования:

- Число рабочих потоков в runtime-системе: 1.
- Кластер НГУ (HP BL2x220c G7, 12 ядер, 24 ГБ ОЗУ на каждом узле).

Задача блочного умножения матриц.

Количество узлов	Централизованный модуль б.н.					Децентрализованный модуль б.н.		
	Время, сек	Ускорение	Эффективность	Параметр1	Параметр2	Время, сек	Ускорение	Эффективность
1	-	-	-	-	-	**60.43**	1.00	100.00
2	63.50	0.95	47.58	1.00	0.50	**38.08**	1.59	79.35
3	31.18	1.94	64.60	0.50	0.10	**29.27**	2.06	68.82
3	31.70	1.91	63.54	1.00	0.50	**29.27**	2.06	68.82
4	**21.25**	2.84	71.09	0.50	0.10	39.93	1.51	37.83
4	**21.81**	2.77	69.27	1.00	0.50	39.93	1.51	37.83
8	**15.16**	3.99	49.83	0.50	0.10	40.72	1.48	18.55
8	**15.71**	3.85	48.08	1.00	0.50	40.72	1.48	18.55
12	**10.12**	5.97	49.76	0.50	0.10	39.66	1.52	12.70
12	**11.71**	5.16	43.00	1.00	0.50	39.66	1.52	12.70
16	**6.72**	8.99	56.20	0.50	0.10	40.69	1.49	9.28
16	**7.73**	7.82	48.86	1.00	0.50	40.69	1.49	9.28

Тестирование

Задача приведения матрицы к диагональному виду.

ФВ либо слишком быстро выполняются, либо зависят от относительно большого размера входных данных по сравнению с объемом вычислений.

Количество узлов	Централизованный модуль б.н.			Децентрализованный модуль б.н.		
	Время, сек	Ускорение	Эффективность	Время, сек	Ускорение	Эффективность
1	-	-	-	6.19	1.00	100.00
2	6.32	0.98	48.97	6.59	0.94	46.97
3	6.41	0.97	32.19	8.01	0.77	25.76
4	6.36	0.97	24.33	8.05	0.77	19.22
8	6.69	0.93	11.57	6.23	0.99	12.42
12	6.73	0.92	7.66	6.68	0.93	7.72
16	6.76	0.92	5.72	6.85	0.90	5.65

Вывод: использование динамического балансировщика нагрузки не имеет смысла для тех программ, в которых размер данных делает время передачи фрагментов вычислений на другой узел сравнимым с временем их выполнения.

Выводы

Разработанное ПО: Система LuNA была расширена модулем централизованной динамической балансировки нагрузки. Модуль был протестирован и показал, что для подходящих программ время их выполнения может значительно сократиться.

Научная новизна: разработан алгоритм централизованной динамической балансировки нагрузки для фрагментированных программ.

Практическая ценность: использование балансировщика может в несколько раз сократить время работы программ, написанных на языке LuNA.

Перспективность: в будущем имеется возможность улучшить способ определения весов ФВ и составление плана балансировки, а также уменьшить накладные расходы.

Публикации

1. Мустафин Д. Э. Модуль централизованной динамической балансировки нагрузки LuNA-программ // Инновации. Наука. Образование. 2021. № 40. С. 365-373.
2. Власенко А. Ю., Мустафин Д. Э., Мичуров М. А. Автоматизация отладки и оптимизации фрагментированных программ в системе LuNA // (статья была принята к опубликованию и будет напечатана в №3 журнала “Проблемы информатики” за 2022 год).
3. Мустафин Д. Э. Реализация централизованного подхода к динамической балансировке вычислительной нагрузки в системе фрагментированного программирования LuNA и его сравнение с децентрализованным подходом // Информационные технологии: Материалы 60-й Междунар. науч. студ. конф. 10–20 апреля 2022 г. / Новосиб. гос. ун-т. — Новосибирск: ИПЦ НГУ, 2022.

Апробация

1. МНСК-2022 // Секция “Параллельные вычисления”.
2. Всероссийская летняя XXXVII молодежная Школа-конференция по параллельному программированию.

Спасибо за внимание!

Планы на будущее

- усовершенствование составления плана балансировки для лучшего удовлетворения приведенных в работе критериев;
- теоретическая оценка накладных расходов на балансировку нагрузки с целью оптимизации параметров, влияющих на общее время работы фрагментированной программы;
- теоретическое сравнение времени исполнения программы с балансировкой централизованной, децентрализованной и без балансировки.

Проверка на необходимость в перераспределении нагрузки

Основные шаги текущего алгоритма балансировки:

1) Если суммарная нагрузка на всех узлах меньше, чем JOBS_LEFT_THRESHOLD, то алгоритм заканчивает свою работу;
2) Находится наименее нагруженный процесс и наиболее нагруженный процесс и разность нагрузки между ними;
3) Если разность нагрузки между двумя процессами меньше, чем JOBS_DIFFERENCE_RATIO, то нет необходимости в перебалансировке, алгоритм заканчивает свою работу;
4) Отправляется сообщение нагруженному процессу о передаче заданного количества ФВ наименее нагруженному процессу;
5) Переход к шагу 2.

ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ГЕОФИЗИКИ СО РАН

ИВМ и МГ СО РАН
пр. Лаврентьева 6
630090, Новосибирск,
РОССИЯ

ПРОБЛЕМЫ ИНФОРМАТИКИ

Тел: +7-383-330-96-43
Mail to: problem-info@sscc.ru
http: problem-info.sscc.ru

СПРАВКА

Настоящим уведомляю, что статья А.Ю. Власенко, М.А. Мичурова, Д.Э. Мустафина «Автоматизация отладки и балансировки нагрузки во фрагментированных программах» была принята к опубликованию и будет напечатана в № 3 журнала «Проблемы информатики» за 2022 год.

Главный редактор
журнала «Проблемы информатики»,
Д. т. н., профессор

В. Э. Малышкин

Для документов
(ИВМиМГ СО РАН)