

Тема работы: Разработка и реализация программных средств автоматизированного обнаружения семантических ошибок во фрагментированных программах для системы LuNA

Состав коллектива:

Мичуров Михаил Антонович, m.michurov@g.nsu.ru, студент ФИТ НГУ

Власенко Андрей Юрьевич, к.т.н., доц. каф. ПВ ФИТ НГУ

Выпускная квалификационная работа

Научный руководитель: Власенко А.Ю., к.т.н., доц. каф. ПВ ФИТ НГУ

Тема утверждена распоряжением проректора по учебной работе от 20 янв 2022 № 0012

Научное содержание работы:

1. Постановка задачи.

Целью данной работы является разработка модуля сбора трассы для системы LuNA и программного средства анализа трассы с целью обнаружения семантических ошибок. Для достижения поставленной цели были поставлены задачи: провести обзор и анализ существующих методов отладки параллельных программ; выделить классы семантических ошибок в LuNA-программах; сформулировать требования к разрабатываемому программному средству; разработать и реализовать обнаружение наиболее типичных для фрагментированных программ семантических ошибок; протестировать разработанное средство на ряде фрагментированных программ, оценить накладные расходы.

2. Современное состояние проблемы.

Разработка параллельных программ численного моделирования на суперЭВМ требует решения комплекса проблем, связанных с организацией параллельной обработки данных и требующих специальной квалификации и больших трудозатрат. Специалисты, в работе которых требуется выполнение численного моделирования, могут испытывать трудности с написанием эффективных параллельных программ.

Для решения этих проблем в ИВМ и МГ СО РАН разрабатывается технология фрагментированного программирования, а также экспериментальная система LuNA (Language for Numeric Algorithms) автоматического конструирования параллельных программ на мультикомпьютерах, поддерживающая данную технологию [1]. Технология фрагментированного программирования представляется перспективной и актуальной для высокопроизводительных вычислений [2].

Система LuNA позволяет автоматически конструировать и выполнять на мультикомпьютерах параллельные программы численного моделирования на основе высокоуровневой спецификации программы на одноименном языке. Однако при написании программ для системы LuNA программист может допускать семантические ошибки [3], не свойственные программам, написанным с использованием традиционных средств параллельного программирования, таких как MPI и OpenMP. Таким образом, проблема отладки LuNA-программ является актуальной.

Для упрощения отладки параллельных программ используются специальные инструментальные средства, реализующие различные подходы к обнаружению ошибок [4]. В рамках данной работы разрабатывалось средство автоматизированной отладки, реализующее подход анализа трассы.

1. Perepelkin V. A. LuNA system for automatic construction of numerical parallel programs for multicomputers // Проблемы информатики. 2020. №1 (46).
2. Маркова В. П., Остапкевич М. Б. Сравнение возможностей MPI и LuNA на примере реализации модели клеточно-автоматной интерференции волн // Проблемы информатики. – 2017. – №. 2 (35). – С. 53-64.

3. Мичуров М. А. Средство анализа причин зависаний фрагментированных программ в системе LuNA // Инновации. Наука. Образование. 2021. № 40. С. 354-364.
4. Власенко А. Ю. и др. Автоматизированный контроль корректности MPI-программ на основе шаблонов ошибочного поведения. – 2014.

3. Подробное описание работы, включая используемые алгоритмы.

Исполнительная система LuNA была расширена модулем, ответственным за сбор трассы исполнения на каждом узле. Этот модуль предоставляет интерфейс для уведомления его о различных событиях исполнительской системы (создание и удаление ФВ и ФД, передача ФВ между узлами и т. п.). Также собранная им информация используется для обнаружения некоторых семантических ошибок во время работы программы, с возможностью последующего подробного анализа после завершения работы.

Утилита анализа трасс, названная `luna_trace`, способна агрегировать трассы, собранные на узлах мультимониторного компьютера во время работы программы, и производить их анализ для выявления семантических ошибок. Утилита предоставляет информацию не только о факте наличия ошибок как таковых, но, по возможности, и о причинах возникновения ошибок с привязкой к исходным кодам LuNA-программы, а также о возможных способах исправления.

4. Полученные результаты.

В результате выполнения работы был проведен анализ семантических ошибок, свойственных фрагментированным программам. Для системы LuNA был разработан модуль сбора трассы, а также реализована возможность автоматической остановки при зависании, вызванном использованием неинициализированных данных. Было разработано средство анализа трассы, выполняющее автоматический поиск семантических ошибок. Разработанное средство протестировано на ряде некорректных LuNA-программ.

Выполнено тестирование с целью оценки влияния сбора отладочной информации на время работы программы, а также с целью оценки времени работы средства анализа трассы в зависимости от параметров задачи.

Разработанное средство в будущем может быть расширено функциональностью по обнаружению дополнительных классов семантических ошибок, а также другими методами и подходами к отладке.

5. Эффект от использования кластера в достижении целей работы.

Кластер использовался для тестирования с целью оценки накладных расходов на сбор трассы в системе LuNA.

Перечень публикаций, содержащих результаты работы

1. Мичуров М. А. Средство анализа причин зависаний фрагментированных программ в системе LuNA // Инновации. Наука. Образование. 2021. № 40. С. 354-364.
2. Власенко А. Ю., Мустафин Д. Э., Мичуров М. А. Автоматизация отладки и оптимизации фрагментированных программ в системе LuNA // (статья принята к опубликованию и будет напечатана в №3 журнала “Проблемы информатики” за 2022 год)
3. Мичуров М. А. Автоматизация обнаружения семантических ошибок во фрагментированных программах для системы LuNA // Информационные технологии: Материалы 60-й Междунар. науч. студ. конф. 10–20 апреля 2022 г.