

ОТЗЫВ

**официального оппонента Галатенко Владимира Антоновича
на диссертационную работу Гимпельсона Вадима Дмитриевича
«Сокращение длины критических путей при динамической
трансляции двоичных кодов»,
представленную к защите на соискание ученой степени кандидата
физико-математических наук по специальности 05.13.11 –
«Математическое и программное обеспечение вычислительных
машин, комплексов и компьютерных сетей»**

Актуальность

Двоичная трансляция является эффективным и широко используемым методом для обеспечения совместимости новых микропроцессоров с программным обеспечением, разработанным ранее для других микропроцессорных архитектур. В отличие от других методов, двоичный транслятор позволяет перенести на новую архитектуру даже то программное обеспечение, для которого недоступны исходные коды. Применение оптимизаций, сокращающих длину критических путей, важно для повышения производительности кода для микропроцессоров с архитектурой широкого командного слова, таких как микропроцессор «Эльбрус». По ряду причин в двоичном трансляторе не могут быть применены оптимизации сокращения длины критических путей, разработанные для компиляторов с языков высокого уровня. Одной из основных причин является жёсткое ограничение на собственное время работы алгоритмов оптимизаций, так как трансляция кодов происходит во время их исполнения. Таким образом, задача разработки быстрых алгоритмов сокращения длины критических путей, которые могут быть использованы в двоичном трансляторе, является **актуальной**.

В работе В.Д. Гимпельсона предложены быстрые алгоритмы сокращения длины критических путей для ациклических и циклических областей. Приведены обоснования корректности и оптимальности алгоритмов, а также оценок их алгоритмической сложности. Обоснования указанных свойств для большинства алгоритмов сформулированы в виде

теорем с соответствующими доказательствами. Предложенные алгоритмы оптимизаций реализованы в двоичном трансляторе для микропроцессора «Эльбрус», что позволило существенно повысить производительность результирующего программного кода.

Структура работы

Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 113 наименований и двух приложений.

Во введении обосновывается актуальность работы, определяются цели и задачи работы, формулируются основные научные результаты, обосновываются теоретическая и практическая ценность, а также приводятся данные об апробации работы.

В первой главе приводится обзор технологии двоичной трансляции как метода обеспечения совместимости новых архитектур микропроцессоров с программным обеспечением, разработанным для старых архитектур, путем исполнения кодов старой архитектуры на новой целевой архитектуре. Приводится также описание EPIC архитектур и особенностей двоичных трансляторов для этого класса архитектур. На основе анализа внутреннего представления программы в двоичном трансляторе показано, что при трансляции двоичного кода возникают множественные необязательные зависимости между операциями. Устранение таких зависимостей позволяет ускорить работу результирующего кода. В конце главы формулируется задача разработки быстрых алгоритмов сокращения длины критических путей в двоичных программах для архитектуры «Эльбрус».

Во второй главе предложены три метода разрыва зависимостей в ациклических областях без построения новых операций: алгоритм переименования регистров, алгоритм применения спекулятивности и использование частичных предикатов. Показано, что данные методы обеспечивают лишь незначительный прирост производительности: 1-2%, при замедлении времени работы транслятора на 0,4%. Тем не менее, рассмотрение и реализация указанных методов важны с точки зрения целостности и полноты диссертационной работы.

В третьей главе рассмотрены методы разрыва зависимостей в ациклических областях с построением новых операций и предложен алгоритм минимизации высоты графа зависимостей. Алгоритм использует техники разрыва зависимостей для антитезисов, зависимостей по результату, предикатных зависимостей, зависимостей между обращениями в память. Суть алгоритма состоит в том, что вначале разрываются все зависимости, которые могут быть разорваны, а затем восстанавливаются те разрывы, которые не уменьшили критический путь. Алгоритм основывается на анализе времен раннего и позднего планирования операций. Приводится формулировка и доказательство теоремы о том, что указанный алгоритм преобразует произвольный граф зависимостей в граф зависимостей с минимально возможной высотой для данных классов преобразований.

Реализация алгоритма в двоичном трансляторе для микропроцессора «Эльбрус» позволила значительно повысить эффективность генерируемого кода. Применение алгоритма на целочисленных и вещественных задачах из пакета SPEC CPU2000 увеличило производительность результирующего кода на 12% и 22% соответственно. На горячих участках Windows 2000 и ряда пользовательских приложений увеличение скорости работы результирующего кода составило 23%. Время работы алгоритма составило 3.8% от общего времени трансляции.

В четвертой главе предложен алгоритм конвейеризации циклов для сокращения критических путей в программных циклах. Идея алгоритма заключается в последовательном переносе через начало цикла некоторых групп операций. Выбор группы операций для переноса на очередном шаге, основывается на разметке времен раннего и позднего планирования на расширенном графе зависимостей. На каждом шаге определяется набор операций, таких что их перенос обеспечивает повышение производительности результирующего кода. Автором разработан алгоритм разметки времен раннего и позднего планирования на расширенном графе зависимостей и доказана его корректность. Разработан быстрый алгоритм конвейеризации циклов и произведена оценка его сложности.

Результаты тестирования показали, что реализация предложенного алгоритма конвейеризации циклов в двоичном трансляторе для микропроцессора «Эльбрус» обеспечила значительный прирост производительности генерируемого кода. На целочисленных и

вещественных задачах из пакета SPEC CPU2000 применение алгоритма обеспечило увеличения скорости работы результирующего кода на 5% и 28% соответственно. На горячих участках ОС Windows 2000 и пользовательских приложений увеличение скорости работы результирующего кода составило 17%. Время работы алгоритма составило 3.5% и 10% от общего времени трансляции на целочисленных и вещественных задачах соответственно.

В заключении подводятся итоги проделанной работы и приводятся основные результаты работы, выносимые на защиту. Приводятся данные, подтверждающие высокую эффективность совокупности разработанных автором методов и алгоритмических решений: в результате их применения в двоичном трансляторе для микропроцессора «Эльбрус» повышение производительности результирующего кода составило, соответственно, 19% и 58% на целочисленных и вещественных задачах из пакета SPEC CPU2000.

Научная новизна

В работе представлены следующие новые научные результаты:

- быстрый алгоритм сокращения длины критического пути в ациклических областях;
- схема взаимодействия алгоритма сокращения длины критического пути в ациклических областях с другими оптимизациями;
- алгоритм разметки времён раннего и позднего планирования на расширенном графе зависимостей;
- быстрый алгоритм конвейеризации циклов;
- интеграция различных методов разрыва зависимостей с алгоритмом конвейеризации циклов;
- сформулированы и доказаны теоремы о корректности и оптимальности предложенных алгоритмов;
- произведена теоретическая и практическая оценка сложности предложенных алгоритмов.

Практическая значимость

Предложенные алгоритмы оптимизаций были реализованы в двоичном трансляторе из архитектуры Intel x86 в архитектуру «Эльбрус», что позволило существенно повысить производительность. Предложенные в данной работе алгоритмы могут быть использованы в двоичных трансляторах и компиляторах с языков высокого уровня для других архитектур.

Достоверность и обоснованность научных положений и выводов работы

Достоверность результатов работы подтверждается апробацией основных результатов на 9 научных конференциях и семинарах. По теме диссертационной работы автором было опубликовано 13 работ, 5 из которых опубликованы в изданиях из списка ВАК, и получено 1 свидетельство о государственной регистрации программы для ЭВМ.

Замечания

По диссертационной работе имеются следующие замечания:

1. В главе 2 предлагается ряд алгоритмов сокращения длины критических путей без построения новых операций. Обоснование этих методов приводится в тексте работы, но не сформулировано в виде соответствующих утверждений или теорем. Формулировка соответствующих утверждений и теорем, как это сделано в главах 3 и 4, была бы уместна и в главе 2.
2. В разделе 4.4 приводится объёмное описание алгоритма разметки времён раннего и позднего планирования на расширенном графе зависимости. Наличие примеров и иллюстраций в описании этого алгоритма упростило бы понимание его сути.
3. Раздел 4.6 посвящён улучшению аппаратной поддержки вращающихся регистров для целей двоичного транслятора. При этом в тексте работы отсутствует описание понятия вращающихся регистров (или, точнее, вращающегося

регистрового файла).

4. В приложениях А и Б приводятся обзоры известных методов конвейеризации циклов, которые могли бы быть заменены ссылками на соответствующие источники.
5. Имеются дефекты в оформлении работы: орфографические и синтаксические (их особенно много) ошибки, опечатки (см., например, с. 17, 42, 94, 144 диссертации, 7, 9, 10 автореферата).

Приведенные замечания не влияют на общую положительную оценку работы.

Заключение

Диссертационная работа соответствует всем требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук, а ее автор, Гимпельсон Вадим Дмитриевич, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.11 – «математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей». Автореферат соответствует основному содержанию диссертационной работы.

Доктор физико-математических наук, заведующий сектором Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

В.А. Галатенко

Пс
На

«26» марта 2018 г.